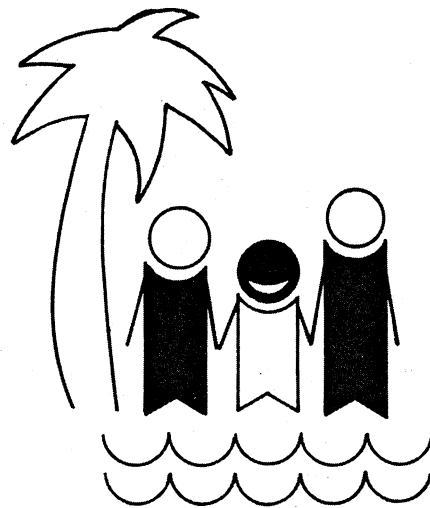


# ***ADVANCES IN CONSUMER RESEARCH***

***VOLUME VI*** <sub>1979</sub>

*Proceedings of the Association  
for Consumer Research  
Ninth Annual Conference  
Miami Beach, Florida  
October 1978*



***WILLIAM L. WILKIE, EDITOR***

1979 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: 0-915552-04-3

International Standard Serial Number: 0098-9258

Wilkie, William L., Editor. *Advances in Consumer Research, Volume VI.* (Ann Arbor: Association for Consumer Research, 1979).

ORDER FORMS ARE PROVIDED AT THE BACK OF THIS VOLUME.

ASSOCIATION FOR CONSUMER RESEARCH - NINTH ANNUAL CONFERENCE

October 26 - 29, 1978 - Miami Beach, Florida

William L. Wilkie, Conference Chairman  
(University of Florida)

H. Keith Hunt, Arrangements Chairman  
(Brigham Young University)

PROGRAM COMMITTEE:

Alden G. Clayton  
Harold H. Kassarian  
Jerome B. Kernan  
Jerry C. Olson  
Peter L. Wright

ACR EXECUTIVE COMMITTEE

Jerome B. Kernan, President  
University of Cincinnati

H. Keith Hunt, President-Elect  
Brigham Young University

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

Thomas C. Kinnear, Executive Secretary  
University of Michigan

Russell A. Bell, Treasurer  
American Greeting Corporation

William L. Wilkie, Conference  
University of Florida

E. Laird Landon, Newsletter  
University of Houston

William B. Locander, Newsletter  
University of Houston

Regina A. Guaraldi, Membership  
Coca-Cola U. S. A.

Mary Jane Schlinger, Publications  
University of Illinois at Chicago Circle

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

Jerry C. Olson, Director  
Pennsylvania State University

Stanley D. Shores, Jr., Director  
Proctor & Gamble

ACR ADVISORY COUNCIL

Kenneth L. Bernhardt  
Georgia State University

Gwen J. Bymers  
Cornell University

John A. Dimling, Jr.  
National Association of Broadcasters

Robert Ferber  
University of Illinois

Stephen A. Greyser  
Harvard University & Marketing Science Institute

Edward Heiden  
U. S. Consumer Product Safety Commission

James R. Taylor  
University of Michigan

Frank D. Walker  
Walker Research, Inc.

William D. Wells  
Needham, Harper, & Steers, Inc.

ORGANIZERS OF SPECIAL TOPICS AND WORKSHOP SESSIONS

Phillippe Cattin, Univ. of Connecticut  
 Alden G. Clayton, Marketing Science Institute  
 Fred G. Cutler, Kennedy Research  
 Thomas Garman, V.P.I.  
 Rebecca Holman, Pennsylvania State Univ.  
 H. Keith Hunt, Univ. of Utah  
 Harold H. Kassarian, UCLA  
 Charles W. King, Purdue Univ.  
 Michael B. Mazis, FTC  
 Dennis McNeill, FTC

J. J. Persensky, National Bureau of Standards  
 Carol A. Scott, UCLA  
 Jagdish N. Sheth, Univ. of Illinois  
 Fred Van Raiij, Univ. of Tilburg  
 M. Venkatesan, Univ. of Oregon  
 Scott Ward, Harvard Univ.  
 William D. Wells, Needham, Harper & Steers Advertising  
 William L. Wilkie, Univ. of Florida  
 Gerald Zaltman, Univ. of Pittsburgh

JUDGES FOR THE PAPER COMPETITION

Olli T. Ahtola, Univ. of Florida  
 Mark I. Alpert, Univ. of Texas  
 Alan R. Andreasen, Univ. of Illinois  
 Gordon G. Bechtel, Univ. of Florida  
 Neil Beckwith, Univ. of Pennsylvania  
 Russell W. Belk, Univ. of Illinois  
 Kenneth L. Bernhardt, FTC  
 James R. Bettman, UCLA  
 Stewart Bither, Pennsylvania State Univ.  
 Paul N. Bloom, Univ. of Maryland  
 Thomas V. Bonoma, Univ. of Pittsburgh  
 Robert E. Burnkrant, Ohio State Univ.  
 Bobby J. Calder, Northwestern Univ.  
 John D. Claxton, Univ. of British Columbia  
 Joel B. Cohen, Univ. of Florida  
 Raymond Dabbah, Pillsbury Co.  
 William R. Darden, Univ. of Arkansas  
 Harry L. Davis, Univ. of Chicago  
 George S. Day, Univ. of Toronto  
 Ralph L. Day, Indiana Univ.  
 F. Steward DeBruicker, Univ. of Pennsylvania  
 John Dimling, National Association of Broadcasters  
 Ira J. Dolich, U. of Nebraska  
 Luis V. Dominguez, Case-Western Reserve Univ.  
 James F. Engel, Wheaton College  
 Ronald E. Frank, Univ. of Pennsylvania  
 David M. Gardner, Univ. of Illinois  
 James L. Ginter, Ohio State Univ.  
 Marvin Goldberg, McGill Univ.  
 Donald H. Granbois, Indiana Univ.  
 Paul E. Green, Univ. of Pennsylvania  
 George H. Haines, Univ. of Toronto  
 Donald J. Hempel, Univ. of Connecticut  
 Gerald E. Hills, Univ. of Tennessee  
 John A. Howard, Columbia Univ.  
 Joel C. Huber, Duke Univ.  
 G. David Hughes, Univ. of North Carolina  
 H. Keith Hunt, Univ. of Utah  
 Jacob Jacoby, Purdue Univ.  
 Harold H. Kassarian, UCLA  
 Thomas C. Kinnear, Univ. of Michigan  
 Donald Kirchner, California State Univ., Northridge  
 Zarrel V. Lambert, Auburn Univ.  
 Clark Leavitt, Ohio State Univ.  
 Donald R. Lehmann, Columbia Univ.  
 V. Parker Lessig, Univ. of Kansas  
 William B. Locander, Univ. of Houston  
 Richard J. Lutz, UCLA  
 J. Barry Mason, Univ. of Alabama  
 Frederick E. May, Univ. of Missouri, St. Louis  
 E. Scott Maynes, Cornell Univ.

Michael B. Mazis, FTC  
 Andrew A. Mitchell, Carnegie-Mellon Univ.  
 Robert Mittelstaedt, Univ. of Nebraska  
 Reza Moinpour, Univ. of Washington  
 Kent B. Monroe, V.P.I.  
 William Moran, Admar Research Company  
 Earl W. Morris, Iowa State Univ.  
 James H. Myers, Univ. of Southern California  
 Joseph W. Newman, Univ. of Arizona  
 Franco Nicosia, University of California, Berkeley  
 Terrence V. O'Brien, Kansas State Univ.  
 Richard W. Olshavsky, Indiana Univ.  
 Sadaomi Oshikawa, Univ. of Washington  
 John W. Payne, Duke Univ.  
 Larry H. Percy, Creamer Lois FSR Inc.  
 Steven Permut, Yale Univ.  
 William D. Perreault, Univ. of North Carolina  
 Edgar A. Pessemier, Purdue Univ.  
 Robert Peterson, Univ. of Texas at Austin  
 Ivan L. Preston, Univ. of Wisconsin  
 Vithala R. Rao, Cornell Univ.  
 Michael L. Ray, Stanford Univ.  
 Richard C. Reizenstein, Univ. of Tennessee  
 Fred D. Reynolds, Univ. of Georgia  
 Thomas S. Robertson, Univ. of Pennsylvania  
 Ivan Ross, Univ. of Minnesota  
 E. Jay Russo, Univ. of Chicago  
 Michael J. Ryan, Columbia Univ.  
 Alan G. Sawyer, Ohio State Univ.  
 David Schoenfeld, J.C. Penney Co.  
 Howard E. Schutz, Univ. of California, Davis  
 Carol A. Scott, UCLA  
 Jagdish N. Sheth, Univ. of Illinois  
 Terrence A. Shimp, Univ. of South Carolina  
 Stanley D. Shores, Proctor & Gamble Co.  
 Alvin J. Silk, MIT  
 Richard Staelin, Carnegie-Mellon  
 Frederick D. Sturdivant, Ohio State Univ.  
 John O. Summers, Indiana Univ.  
 John E. Swan, Univ. of Alabama, Birmingham  
 James R. Taylor, Univ. of Michigan  
 Douglas J. Tigert, Univ. of Toronto  
 Frank D. Walker, Walker Research  
 Scott Ward, Harvard Univ.  
 John J. Wheatly, Univ. of Washington  
 Albert R. Wildt, Univ. of Florida  
 James B. Wiley, Temple Univ.  
 William L. Wilkie, Univ. of Florida  
 Arch Woodside, Univ. of South Carolina  
 Peter L. Wright, Stanford Univ.  
 Gerald Zaltman, Univ. of Pittsburgh

## PREFACE

The Association for Consumer Research (ACR) was founded for the purpose of fostering the development of a fledgling field; a commitment to quality has been one of its primary characteristics. Each year the Association has grown along with the field of consumer behavior. The Advances in Consumer Research series is a chronicle for this development.

Volume VI of the series presents most of the substance of ACR's Ninth Annual Conference, held at the Carillon Hotel, Miami Beach, October 26-29, 1978. All 81 of the papers selected in the paper competition are included in this volume, as are summaries of most of the Workshop and Special Topics sessions, and the Presidential Address by Dr. Jerome B. Kernan.

In addition, a special feature of this volume is the set of 45 commentaries by leading persons in the field of consumer behavior. These "discussion papers" were prepared specifically by topic, with 2 papers per session. In addition to commenting upon the strengths and limitations of the session's papers, most authors go on to address the topic more generally, proposing significant issues and avenues for future research. One author, Edward C. Strong, goes so far as to present a literary contribution.

A number of people made substantial contributions to both the Conference and to this Proceedings. The key contributions are those of the paper authors and the organizers of the special topic and workshop sessions. The lasting value of the conference is clearly due to their individual efforts. In addition, a group of 95 judges played a special role in determining the substance of the meetings, and provided useful feedback to many persons in this field. All of these groups are listed elsewhere in this Volume. My special thanks must go to Professor H. Keith Hunt, who in addition to handling arrangements for the Conference, often provided me with much needed advice and suggestions.

Closer to home, several persons at the University of Florida made substantial contributions. Peter R. Dickson, Ph.D. Candidate, went far beyond his assigned duties, particularly in providing a much needed structure for administering the Conference. Melanie Albert has picked up where Peter left off, and has been instrumental in compiling these Proceedings. In addition, Mary Ellen Adams, Harriet High and Nancy C. Peat volunteered to proof the paper admissions here.

Thanks go to the staff of the Marketing Department who carried out most of the work. Mrs. Lois Wiechmann coordinated the effort, and together with Vicki Avery handled the typing and correspondence. They were ably assisted by Janet Fitzpatrick and Vada Taylor. Finally, the provision of some released time was essential in allowing me to undertake this activity. My appreciation is extended to Joel B. Cohen and the University of Florida Marketing Department for supporting this effort.

At home, it's been a hectic time, punctuated by the simultaneous arrival of Billy and the closing date for Conference planning. My wife Barbara bore the brunt of this. It doesn't seem appropriate to say anything more in public, but I will do so privately.

Gainesville, Florida  
February, 1979

William L. Wilkie  
Editor

*"Conflict is the gadfly of thought. It stirs us to observation and memory. It instigates to invention. It shocks us out of sheeplike passivity, and sets us at noting and contriving...conflict is a sine qua non of reflection and ingenuity."*

John Dewey, *Morals and Conduct*

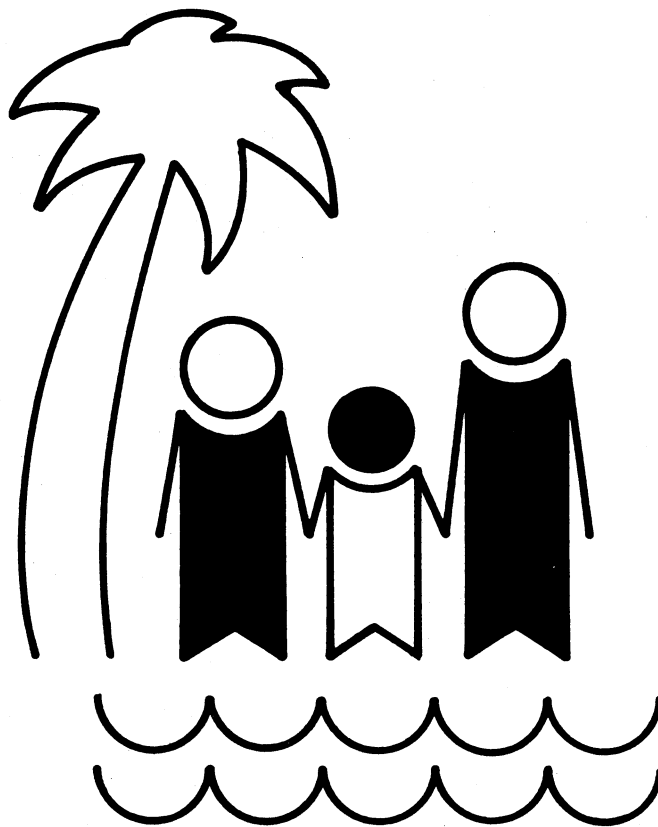
*"Cogito cogito, ergo cogito sum."*

Ambrose Bierce, *The Devil's Dictionary*

(1978 PROGRAM)  
**Association for Consumer Research**  
**Ninth Annual Conference**

**Carillon Hotel**  
**Miami Beach, Florida**

**October 26 - 29, 1978**



**William L. Wilkie, Program Chairman**  
**H. Keith Hunt, Arrangements Chairman**

**Program Committee:**

**Alden G. Clayton**  
**Harold H. Kassarian**  
**Jerome B. Kernan**  
**Jerry C. Olson**  
**Peter L. Wright**

# Program 1978: Annual Conference of the Association For Consumer Research

## Thursday, October 26, 1978

Registration opens at 5:00 p.m. and will remain open until 9:00 p.m.  
 EARLY BIRD RECEPTION 6:00 - 7:30 (Cash Bar)  
 Evening: Dinner on your own. No evening sessions.

## Friday, October 27, 1978

Registration opens at 8:00 a.m. and will remain open until 4:00 p.m.

### FRIDAY MORNING SESSIONS — I

Delmonico East — 9:00 - 11:00

(Workshop)

#### “CAN AND SHOULD THE FTC RESTRICT ADVERTISING TO CHILDREN?”

CHAIRMAN: *Michael B. Mazis*, Federal Trade Commission

#### PANEL MEMBERS:

*Tracy Weston*, Assistant Director, Bureau of Consumer Protection, Federal Trade Commission  
*Seymour Banks*, Vice-President, Leo Burnett Advertising  
*John Dimling*, Vice-President, National Association of Broadcasters  
*Robert Choate*, President, Council on Children, Media and Merchandising  
*Gerald Thain*, Professor, Univ. of Wisconsin school of Law  
*Stanley Cohen*, Washington Editor, *Advertising Age*

Live Oak Room — 9:00 - 10:30

(Workshop)

#### “LONGITUDINAL RESEARCH DESIGNS AND THE ANALYSIS OF PANEL DATA”

CHAIRMAN: *Fred G. Cutler*, Kennedy Research

#### PANEL MEMBERS:

*Lois Benedetti*, ATT&T Long Lines  
*Donald R. Lehmann*, Columbia Univ.  
*Randall Harris*, National Family Opinion, Inc.  
*Lorraine Scarpa*, Heublein, Inc.  
*Arch Woodside*, Univ. of South Carolina

Banyan East — 9:00 - 10:30

#### “INVOLVEMENT”

CHAIRMAN: *David M. Gardner*, Univ. of Illinois

“Questioning the Concept of Involvement Defined Product Classes”

*John L. Lastovicka*, Temple Univ.

“An Examination of Ego-Involvement as a Modifier of Attitude Changes Caused from Product Testing”

*Larry M. Newman* (student), Pennsylvania State Univ.

*Ira J. Dolich*, Univ. of Nebraska

“Exploring Generalized and Personalized Beliefs Among Smokers and Non-Smokers: A First Look”

*Kjell Grønhaug*, The Norwegian School of Economics and Business  
*Norman Kangun*, Univ. of Houston

#### DISCUSSION PAPERS:

*Andrew A. Mitchell*, Carnegie-Mellon Univ.

*Michael L. Ray*, Stanford Univ.

**BREAK 11:00 - 11:15**

**BREAK 10:30 - 10:45**

**BREAK 10:30 - 10:45**

### FRIDAY MORNING SESSIONS — II

Delmonico East — 11:15 - 12:15

(Workshop)

#### “RESEARCHERS LOOK AT THE KID VID RULE”

CHAIRMAN: *Scott Ward*, Harvard Univ.

#### PANEL MEMBERS:

*Charles Atkin*, Michigan State Univ.  
*Marvin Goldberg*, McGill Univ.  
*Thomas Robertson*, Univ. of Pennsylvania  
*Ellen Wartella*, Ohio State Univ.

Banyan East — 10:45 - 12:15

(Special Topic)

#### “TIME: THE FUNDAMENTAL THINGS APPLY...”

CHAIRPERSONS: *M. Venkatesan*, Univ. of Oregon  
*Rebecca Holman*, Pennsylvania State Univ.

#### SPEAKERS:

*Douglas Hawes*, Univ. of Wyoming  
*Phillip Hendrix*, Florida State Univ.  
*Rebecca Holman*, Pennsylvania State Univ.  
*Thomas C. Kinnear*, Univ. of Michigan  
*Robert Settle*, San Diego State Univ.  
*James R. Taylor*, Univ. of Michigan  
*Christine Urban*, Harvard Univ.  
*M. Venkatesan*, Univ. of Oregon

Cypress Room — 10:45 - 12:15

(Workshop)

#### “FURTHER DEBATE ON THE MEASUREMENT OF ATTRIBUTE UTILITIES”

CHAIRMEN: *Phillippe Cattin*, Univ. of Connecticut  
*Dick Wittink*, Stanford Univ.

#### PANEL MEMBERS:

*Gordon Bechtel*, Univ. of Florida  
*Paul Green*, Univ. of Pennsylvania  
*William Moran*, Admar Research, Inc.  
*Martin Starr*, Columbia Univ.  
*Richard Westwood*, David Beaumont Organization

**LUNCHEON 12:15 - 1:30 Delmonico West**

**Research Interest Tables.**



**NOTES**  
9:00 - 10:30

**FRIDAY MORNING SESSIONS – I**

Cypress Room – 9:00 - 10:30

**“PERSONALITY AND CONSUMER TRAITS”**

CHAIRMAN: *Peter D. Bennett*, Pennsylvania State Univ.

“An Exploratory Study of Assertiveness, Aggressiveness, and Consumer Complaining Behavior”

*Claes Fornell*, Northwestern Univ.  
*Robert A. Westbrook*, Univ. of Arizona

“Correlations Between Three Indicators of Breadth and Variation in Sources of Stimulation”

*Elizabeth C. Hirschman*, New York Univ.  
*Melanie Wallendorf*, Univ. of Michigan

“Dogmatism and Innovation: A Situational Perspective”

*Kenneth A. Coney*, Arizona State Univ.  
*Robert R. Harmon*, Portland State Univ.

DISCUSSION PAPERS:

*Harold H. Kassarian*, UCLA  
*Jerome B. Kernan*, Univ. of Cincinnati

Banyan West – 9:00 - 10:30

**“METHODOLOGICAL DEVELOPMENTS”**

CHAIRMAN: *Terrence V. O'Brien*, Kansas state Univ.

“Changes in Consumer Perceptions: The Impact of Testing Conditions on Perceptions of Branded Products”

*James McCullough*, Univ. of Arizona  
*Charlene S. Martinsen*, Univ. of Washington  
*Linda Scurman* (student), Univ. of Washington

“Covariance Bias of Thurstone Case V Scaling as Applied to Consumer Preferences and Purchase Intentions”

*Joel Huber*, Duke Univ.  
*Murphy A. Sewall*, Univ. of Connecticut

“Market Shares Estimates Based on Conjoint Analysis of Concepts”

*James B. Wiley*, Temple Univ.  
*Robert Bushnell*, Wayne State Univ.

DISCUSSION PAPERS:

*Robert C. Blattberg*, Univ. of Chicago  
*James L. Ginter*, Ohio State Univ.

---

**BREAK 10:30 - 10:45**

**BREAK 10:30 - 10:45**

---

**FRIDAY MORNING SESSIONS – II**

Live Oak Room – 10:45 - 12:15

**“DEMOGRAPHICS, SPENDING, AND LEISURE”**

CHAIRMAN: *Fred D. Reynolds*, Univ. of Georgia

“Family Life Cycle As A Determinant of Size and Composition of Household Expenditures”

*Johan Arndt*, Norwegian School of Economics and Business

“Family Life Cycle and Leisure Behavior Research”

*E. Laird Landon, Jr.*, Univ. of Houston  
*William B. Locander*, Univ. of Houston

“Social Class Determinants of Leisure Activity”

*Robert B. Settle*, San Diego State Univ.  
*Pamela L. Alreck*, San Diego State Univ.  
*Michael A. Belch*, San Diego State Univ.

DISCUSSION PAPERS:

*Robert Ferber*, Univ. of Illinois  
*Charles W. King*, Purdue Univ.

Banyan West – 10:45 - 12:15

**“PSYCHOLOGICAL CONTRIBUTIONS”**

CHAIRMAN: *Clark Leavitt*, Ohio State Univ.

“Value Structures and Consumer Behavior”

*Jonathon Gutman*, Univ. of Southern California  
*Donald E. Vinson*, Univ. of Southern California

“Measuring the Bases of Social Power”

*John L. Swasy*, Pennsylvania State Univ.

“Deviant Consumer Behavior: A Different View”

*Michael K. Mills* (student), Univ. of Pittsburgh  
*Thomas V. Bonoma*, Univ. of Pittsburgh

DISCUSSION PAPERS:

*Stewart Bither*, Pennsylvania State Univ.  
*Jagdish N. Sheth*, Univ. of Illinois

**NOTES**  
10:45 - 12:15

**FRIDAY AFTERNOON SESSIONS – III**

Banyan West – 1:30 - 3:00

(Workshop)

**“THE ENERGY CRISIS AND CONSUMER CONSERVATION: CURRENT RESEARCH AND ACTION PROGRAMS”**

CHAIRMAN: *William L. Wilkie*, Univ. of Florida

PANEL MEMBERS:

*Brian Kelly*, Canadian Department of Energy, Mines, & Resources  
*Jeffrey Millstein*, U.S. Department of Energy  
*R. Bruce Hutton*, U.S. Department of Energy

Cypress Room – 1:30 - 3:00

(Workshop)

**“CONSUMER RESEARCH IN ACTION: HOW, WHEN, WHERE, AND WHY RESEARCH IS APPLIED IN DAY TO DAY OPERATIONS IN INDUSTRY”**

CHAIRMAN: *Alden G. Clayton*, Managing Director, Marketing Science Institute

PANEL MEMBERS:

*Thomas C. Dunkerton*, Senior Vice-President, Compton Advertising, Inc.  
*Donald Hughes*, Manager, Consumer Research Division, Sears, Roebuck and Company  
*Lawrence M. Krueger*, Vice-President of Marketing Services, Burger King International  
*Edward W. Hart, Jr.*, Associate Manager Consumer Research, Heublein, Inc.

Banyan East – 1:30 - 3:00

**“KEY QUESTIONS IN VALIDITY”**

CHAIRMAN: *Michael J. Ryan*, Columbia Univ.

“On the Use of Formulas of the Predictive Validity of Regression in Consumer Research”

*Philippe Cattin*, Univ. of Connecticut

“Construct Validation in Marketing: A Comparison of Methods in Assessing the Validity of the Affective, Conative, and Cognitive Components of Attitudes”

*George John* (student), Northwestern Univ.

*Torger Reve* (student), Northwestern Univ.

“Attitude Measurement and Behavior Change: A Reconsideration of Attitude Organization and Its Relationship to Behavior”

*Richard P. Bagozzi*, Univ. of California, Berkeley

*Robert E. Burnkrant*, Ohio State Univ.

DISCUSSION PAPERS:

*Joel B. Cohen*, Univ. of Florida

*Richard Staelin*, Carnegie-Mellon Univ.

---

**BREAK 3:00 - 6:00**

**BREAK 3:00 - 6:00**

**BREAK 3:00 - 6:00**

---

**COCKTAIL HOUR 6:00 - 7:30**

**Dinner on your own. No evening sessions.**

**FRIDAY AFTERNOON SESSIONS — III**

Live Oak Room — 1:30 - 3:00

**“FAMILY DYNAMICS”**

CHAIRMAN: *Donald J. Hempel*, Univ. of Connecticut

“Family Communication and Consumer Socialization”

*George P. Moschis*, Georgia State Univ.  
*Roy L. Moore*, Georgia State Univ.

“A Cross-National Exploration of Husband-Wife Involvement in Selected Household Activities”

*Susan P. Douglas*, Centre d'Enseignement Supérieur des Affaires

“Underlying Perceptual Patterns in Husband and Wife Purchase Decision Influence Assessments”

*Alvin C. Burns*, Louisiana State Univ.  
*David J. Ortinau* (student), Louisiana State Univ.

**DISCUSSION PAPERS:**

*Harry L. Davis*, Univ. of Chicago  
*Donald H. Granbois*, Indiana Univ.

**NOTES**

1:30 - 3:00

Delmonico East — 1:30 - 3:00

**KID'S TV RULE (If continued)**

---

**BREAK 3:00 - 6:00**

**BREAK 3:00 - 6:00**

---

## Saturday, October 28, 1978

Registration open 8:30 a.m. to 5:00 p.m.

### SATURDAY MORNING SESSIONS — I

Cypress Room — 9:00 - 10:30

(Special Topic)

#### “PRAGMATIC APPLICATIONS OF CONSUMER RESEARCH IN RETAILING”

CHAIRMAN: *Charles W. King*, Purdue Univ.

#### SPEAKERS:

*Vern L. Page*, National Merchandise Manager, Sears, Roebuck & Co.  
*Ralph D. Holt*, Research Director, Federated Department Stores  
*Jerry K. Townley*, Group Director/Marketing, Burger King Corp.  
*William Applebaum*, Lecturer on Food Distribution  
*Lawrence J. Ring*, Univ. of Virginia  
*Douglas J. Tigert*, Univ. of Toronto

Banyan West — 9:00 - 10:30

(Workshop)

#### “CONSUMER RESEARCH GOES TO WASHINGTON”

CHAIRMAN: *J. J. Persensky*, National Bureau of Standards

#### PANEL MEMBERS:

*James Bradley*, Consumer Product Safety Commission  
*George E. Brosseau*, Administration on Aging  
*Chuck Handy*, U. S. Department of Agriculture  
*Dennis McNeill*, Federal Trade Commission  
*Louis Morris*, Food and Drug Administration  
*Anne Ramey Smith*, National Bureau of Standards  
*Raymond Stokes*, Food and Drug Administration

Delmonico East — 9:00 - 10:30

#### “CONSUMER DECISION MAKING”

CHAIRMAN: *Stanley D. Shores*, Proctor & Gamble Co.

“Stimulus-Response Variables in New Product Research”

*P. S. Raju*, Pennsylvania State Univ.

“The Effects of Choice Complexity and Decision Freedom on Choice Behavior”

*John R. Walton*, Univ. of Minnesota  
*Eric N. Berkowitz*, Univ. of Minnesota

“Evaluative Conflict and Information Search In The Adoption Process”

*James W. Harvey*, Univ. of Maryland

#### DISCUSSION PAPERS:

*James R. Bettman*, UCLA  
*Peter L. Wright*, Stanford Univ.

---

BREAK 10:30 - 10:45

BREAK 10:30 - 10:45

BREAK 10:30 - 10:45

### SATURDAY MORNING SESSIONS — II

Live Oak Room — 10:45 - 12:15

(Special Topic)

#### “RESEARCH ADVANCES IN HOME ECONOMICS: WHY GIVE A DAMN?”

CHAIRMAN: *E. Thomas Garman*, Professor, Dept. of Management, Housing and Family Development, V.P.I.

#### SPEAKERS:

*Patricia Sailor*, Director, School of Home Economics, Louisiana State Univ.  
*Lucille Wakefield*, Head, Dept. of Food and Nutrition, Florida State Univ.  
*E. Scott Maynes*, Chairman, Dept. of Consumer Economics, Cornell Univ.  
*Earl W. Morris*, Professor, Dept. of Family Environment, Iowa State Univ.

Banyan West — 10:45 - 12:15

#### “BACK TO THE BASICS”

CHAIRMAN: *David T. Wilson*, Pennsylvania State Univ.

“What Do You Learn Standing in a Supermarket Aisle?”

*K. W. Kendall*, Simon Fraser Univ.  
*Ian Fenwick*, Dalhousie Univ.

“Allison & Uhl Revisited: The Effects of Taste and Brand Name on Perceptions and Preferences”

*G. A. Mauser*, Simon Fraser Univ.

“Pupil Dilation Measures in Consumer Research: Applications and Limitations”

*David C. Arch* (student), UCLA

#### DISCUSSION PAPERS:

*James F. Engel*, Wheaton College  
*Joseph W. Newman*, Univ. of Arizona

Delmonico East — 10:45 - 12:15

#### “FRONTIER ISSUES IN ADVERTISING”

CHAIRMAN: *Larry H. Percy*, Creamer Lois FSR Inc.

“A Demonstration of Levels-Of-Processing Theory in Memory for Advertisements”

*Joel Saegert*, Univ. of Texas, San Antonio

“Perceptual Discrepancies in the Time Duration and Number of Television Commercials”

*Peter H. Webb*, Univ. of Illinois

“An Experimental Analysis of Attitudes Toward Comparison and Non-Comparison Advertising”

*Edwin C. Hackleman*, Univ. of Connecticut  
*Subhash C. Jain*, Univ. of Connecticut

#### DISCUSSION PAPERS:

*John G. Myers*, Univ. of California, Berkeley  
*Alan G. Sawyer*, Ohio State Univ.

### LUNCHEON BUFFET 12:15 - 2:00 Delmonico West

Presidential Address: *Jerome B. Kernan*

#### Award Presentations

ACR business meeting will commence during the second half of the luncheon. Those not wishing to attend should feel free to leave whenever the urge strikes.

**SATURDAY MORNING SESSIONS – I**

Banyan East – 9:00 - 10:30

**“ATTRIBUTES: IMPORTANT, SALIENT AND IDEAL”**

CHAIRMAN: *Terrence A. Shimp*, Univ. of South Carolina

“A Comparison of Two Methods For Determining Optimum Levels of Product Characteristics”

*James H. Myers*, Univ. of Southern California  
*Richard F. Chay*, American Cyanamid

“Behavioral Measurement of the Relative Importance of Attribute-Related Information Cues: The Case of Cold Breakfast Cereals”

*John A. Quelch*, Univ. of Western Ontario

“The Stability of Responses Obtained By Free Elicitations: Implications For Measuring Attribute Salience and Memory Structure”

*Jerry C. Olson*, Pennsylvania State Univ.  
*Aydin Muderrisoglu* (student), Pennsylvania State Univ.

DISCUSSION PAPERS:

*Richard P. Bagozzi*, Univ. of California, Berkeley  
*Donald R. Lehmann*, Columbia Univ.

---

**BREAK 10:30 - 10:45**

**SATURDAY MORNING SESSIONS – II**

Banyan East – 10:45 - 12:15

**“FURTHER DEBATE ON MULTIATTRIBUTE ATTITUDES”**

CHAIRMAN: *Reza Moïnpour*, Univ. of Washington

“A Comment on the State of Attitude Measurement in Consumer Research: A Polemic”

*E. H. Bonfield*, Temple Univ.

“A Note on the Fishbein Attitude Model as an Expected Utility Model”

*Joel Huber*, Duke Univ.  
*Robert P. Leone*, Univ. of Texas

“Multiattribute Preference Models for Consumer Research: A Synthesis”

*Arun K. Jain*, SUNY, Buffalo  
*Vijay Mahajan*, Ohio State Univ.  
*Naresh K. Malhotra* (student), SUNY, Buffalo

DISCUSSION PAPERS:

*Olli T. Ahtola*, Univ. of Florida  
*Sadaomi Oshikawa*, Univ. of Washington

Live Oak Room – 9:00 - 10:30

**“SIGNIFICANT PROBLEMS IN ESTIMATION”**

CHAIRMAN: *V. Parker Lessig*, Univ. of Kansas

“Estimating the Effects of Advertising: Application To A Social Problem”

*James G. Barnes*, Memorial Univ. of Newfoundland  
*Jacques C. Bourgeois*, Carleton Univ.

“Assessing the Potential Effects of Differential Price Increases on Gasoline Usage”

*Richard C. Reizenstein*, Univ. of Tennessee  
*David J. Barnaby*, Univ. of Tennessee

“Drug Compliance And The Neglected Concern For Validity”

*F. Robert Dwyer*, Univ. of Cincinnati

DISCUSSION PAPERS:

*Neil E. Beckwith*, Univ. of Pennsylvania  
*Albert R. Wildt*, Univ. of Georgia

---

**BREAK 10:30 - 10:45**

Cypress Room – 10:45 - 12:15

**“THE CONSUMER ENVIRONMENT: REGULATION AND RESEARCH”**

CHAIRMAN: *Gary Ford*, Univ. of Maryland

“Research on Consumer Information: Public Sector Perspectives”

*Steven E. Permut*, Yale Univ.

“Consumer Research Inputs Into Public Policy Decision Making: The Role of Canada’s Consumer Research and Evaluation Branch”

*John L. Evans*, Consumer & Corporate Affairs Canada  
*C. Dennis Anderson*, Univ. of Manitoba  
*L. G. McCabe*, Consumer & Corporate Affairs Canada

“Consumer Problems and Complaints: A National View”

*Marc A. Grainer*, Tech. Assistance Research Programs  
*Kathleen A. McEvoy*, King Research  
*Donald W. King*, King Research

DISCUSSION PAPERS:

*Stephen A. Greyser*, Harvard Univ. and Marketing Science Inst.  
*Ivan L. Preston*, Univ. of Wisconsin

**NOTES**  
9:00 - 10:30

**NOTES**  
10:45 - 12:15

**SATURDAY AFTERNOON SESSIONS – III**

Banyan East – 2:00 - 4:00

(Workshop)

**“THE ORIGINS AND FUTURE OF ACR”**

CHAIRMAN: *Harold H. Kassarian, UCLA*

SPEAKER:

*James F. Engel, Wheaton College*

PANEL OF ACR PRESIDENTS:

*Joel B. Cohen, Univ. of Florida*  
*Robert W. Pratt, Heublein, Inc.*  
*William D. Wells, Needham, Harper, & Steers, Inc.*  
*Jacob Jacoby, Purdue Univ.*  
*David M. Gardner, Univ. of Illinois*  
*Jerome B. Kernan, Univ. of Cincinnati*

SUMMARY PRESENTATION of JCR  
Award-winning paper by author.

Cypress Room – 2:00 - 3:30

**“RESEARCH CONTRIBUTIONS TO PUBLIC POLICY”**

CHAIRMAN: *Kenneth L. Bernhardt, Federal Trade Commission*

“A Study of Public Policy Impact on Consumer Decision-Making”

*Philip C. Burger, SUNY, Binghamton*  
*Alladi Venkatesh, SUNY, Binghamton*

“A Measure for Market Delineation”

*Jacques C. Bourgeois, Carleton Univ.*

“Attracting Potential Switchers to Mass Transit: Mode Choice as a Multiattribute Decision Model”

*Linda L. Golden, Univ. of Texas*  
*John F. Betak, Univ. of Texas*  
*Mark I. Alpert, Univ. of Texas*

DISCUSSION PAPERS:

*Ivan Ross, Univ. of Minnesota*  
*Frederick D. Sturdivant, Ohio State Univ.*

Live Oak Room – 2:00 - 3:30

**“INFORMATION PROCESSING AND EVOKED SETS”**

CHAIRMAN: *F. Stewart DeBruicker, Univ. of Pennsylvania*

“Evoked Set Formation and Composition: The Learning and Information Processing Hypotheses”

*Frederick E. May, Univ. of Missouri, St. Louis*

“An Information Processing Approach to Evoked Set Formation”

*Thomas L. Parkinson, Pennsylvania State Univ.*  
*Michael Reilly, Pennsylvania State Univ.*

“Decision Rule Uncertainty, Evoked Set Size and Task Difficulty as a Function of Number Choice Criteria and Information Variability”

*Joseph J. Belonax, Jr., Western Michigan Univ.*

DISCUSSION PAPERS:

*Bobby J. Calder, Northwestern Univ.*  
*James H. Myers, Univ. of Southern California*

---

**BREAK 4:00 - 4:15**

**BREAK 3:30 - 3:45**

**BREAK 3:30 - 3:45**

---

**SATURDAY AFTERNOON SESSIONS – IV**

Banyan East – 4:15 - 5:15

(Special Topic)

**“THE ESSENTIALS AND PROMISE OF LIFESTYLE RESEARCH”**

CHAIRMAN & SPEAKER:

*William D. Wells, Needham, Harper & Steers Advertising*  
(plus video presentation)

Banyan West – 3:45 - 5:15

(Special Topic)

**“STRATEGIES OF PLANNED CHANGE”**

CO-CHAIRMEN: *Jagdish N. Sheth, Univ. of Illinois*  
*Gerald Zaltman, Univ. of Pittsburgh*

SPEAKERS:

*Christopher Lovelock, Harvard Univ.*  
*Steven E. Permut, Yale Univ.*  
*Robert Rich, Princeton Univ.*  
*Jagdish Sheth, Univ. of Illinois*  
*Gerald Zaltman, Univ. of Pittsburgh*

Cypress Room – 3:45 - 5:15

**“SITUATIONAL IMPACTS ON CONSUMER PURCHASING: GIFT-GIVING”**

CHAIRMAN: *Kent B. Monroe, V.P.I.*

“The Effects of Product Involvement and Task Definition on Anticipated Consumer Effect”

*Thomas K. Clarke (student), Univ. of Illinois*  
*Russell W. Belk, Univ. of Illinois*

“Gift-Giving: A Review and an Interactive Paradigm”

*Sharon K. Banks (student), Univ. of Oregon*

“Gift Versus Personal Use Brand Selection”

*Roger Heeler, York Univ.*  
*June Francis (student), York Univ.*  
*Chike Okechuku (student), York Univ.*  
*Stanley Reid (student), York Univ.*

DISCUSSION PAPERS:

*Richard J. Lutz, UCLA*  
*Douglas J. Tigert, Univ. of Toronto*

**COCKTAIL HOUR 6:00 - 7:30**

Dinner on your own. Enjoy!

**SATURDAY AFTERNOON SESSIONS — III**

Delmonico East — 2:00 - 3:30

**“ISSUES IN SURVEY MEASUREMENT”**

CHAIRMAN: *Michael Houston*, Univ. of Wisconsin

“Survey Data Reliability Effects on Results of Consumer Preference Analyses”

*Abraham D. Horowitz*, General Motors Research Labs  
*Thomas F. Golob*, General Motors Research Labs

“The Effect of Varying Response Intervals on the Stability of Factor Solutions of Rating Scale Data”

*Roger Best*, Univ. of Arizona  
*Del I. Hawkins*, Univ. of Oregon  
*Gerald Albaum*, Univ. of Oregon

“The Effects of Sponsorship on Mail Survey Response and Evaluation Bias”

*Kenneth E. Miller*, Univ. of Utah  
*Marilyn Turner* (student), Univ. of British Columbia

DISCUSSION PAPERS:

*Gordon G. Bechtel*, Univ. of Florida  
*Paul E. Green*, Univ. of Pennsylvania

**NOTES**

2:00 - 3:30

Banyan West — 2:00 - 3:30

**“SOCIOLOGICAL PERSPECTIVES”**

CHAIRMAN: *William R. Darden*, Univ. of Arkansas

“Extending Product Life: Technology Isn't the Only Issue”

*Margaret DeBell* (student), Univ. of Maryland  
*Rachel Dardis*, Univ. of Maryland

“A Path-Analytic Investigation of Life Satisfaction Among Elderly Consumers”

*William O. Bearden*, Univ. of South Carolina  
*A. William Gustafson*, Texas Tech. Univ.  
*J. Barry Mason*, Univ. of Alabama

“Smoking Behaviors as a Diffusion Process Within Age Cohort Groups: An Application of the Societal Marketing Concept”

*Theodore F. Smith*, Old Dominion Univ.  
*Richard W. Olshavsky*, Indiana Univ.  
*Michael F. Smith* (student), Indiana Univ.

DISCUSSION PAPERS:

*Thomas S. Robertson*, Univ. of Pennsylvania  
*Gerald Zaltman*, Univ. of Pittsburgh

---

**BREAK 3:30 - 3:45**

**BREAK 3:30 - 3:45**

---

**SATURDAY AFTERNOON SESSIONS — IV**

Live Oak Room — 3:45 - 5:15

**“CONSUMER POLICY PROPOSALS”**

CHAIRMAN: *Howard E. Schutz*, Univ. of California, Davis

“Formal Consumer Education: An Empirical Assessment”

*George P. Moschis*, Georgia State Univ.

“On Consumer Dissatisfaction: Consumer Arbitration as an Alternative Dispute Resolution Mechanism”

*Stephen A. Goodwin*, State Univ. of New York at Buffalo  
*Vijay Mahajan*, Ohio State Univ.  
*Bhal J. Bhatt*, State Univ. of New York at Buffalo

“In Search of Actions to Reduce Consumer Shopping Problems”

*J. R. Brent Ritchie*, Univ. of Calgary  
*John D. Claxton*, Univ. of British Columbia

DISCUSSION PAPERS:

*E. Scott Maynes*, Cornell Univ.  
*Hans B. Thorelli*, Indiana Univ.

Delmonico East — 3:45 - 5:15

**“SEX IN ADVERTISING”**

CHAIRPERSON: *Alice Courtney*, York Univ.

“The Role of Sexually-Oriented Stimuli in Advertising: Theory and Literature Review”

*R. Dale Wilson*, Pennsylvania State Univ.  
*Noreen K. Moore* (student), Pennsylvania State Univ.

“Demographic and Cognitive Factors Influencing Viewers' Evaluations of 'Sexy' Advertisements”

*Donald Sciglimpaglia*, San Diego State Univ.  
*Michael A. Belch*, San Diego State Univ.  
*Richard F. Cain, Jr.* (student), San Diego State Univ.

“How Should Women be Portrayed in Advertisements? — A Call for Research”

*Mary Lou Roberts*, Boston Univ.  
*Perri B. Koggan*, Boston Univ.

DISCUSSION PAPERS:

*Mark I. Alpert*, Univ. of Texas  
*Edward C. Strong*, Tulane Univ.

**NOTES**

3:45 - 5:15

**Sunday, October 29, 1978**

**SUNDAY MORNING SESSIONS — I**

Banyan East — 9:00 - 12:15

(Special Topic)

**“WHAT IS ATTRIBUTION THEORY, AND WHAT DOES IT HAVE TO DO WITH CONSUMER BEHAVIOR?”** (3 hours)

CHAIRPERSON: *Carol A. Scott, UCLA*

SPEAKERS:

*Peter Reingen, Univ. of South Carolina*  
*Carol Scott, UCLA*  
*Alice Tybout, Northwestern Univ.*

PANEL MEMBERS:

*Robert Burnkrandt, Ohio State Univ.*  
*Bobby J. Calder, Northwestern Univ.*  
*Richard Mizerski, Univ. of Cincinnati*  
*Robert Smith, Indiana Univ.*  
*Brian Sternthal, Northwestern Univ.*  
*William Swinyard, Brigham Young Univ.*  
*Valerie Valle, Univ. of Pittsburgh*

Live Oak Room — 9:00 - 12:15

(Workshop)

**“CONSUMER RESEARCH IN EUROPE: FRONTIER ISSUES”** (3 hours)

CHAIRMAN: *Fred Van Raaij, Univ. of Tilburg (The Netherlands)*

SPEAKERS:

*Flemming Hansen, Univ. of Copenhagen (Denmark)*  
*Jagdish N. Sheth, Univ. of Illinois*  
*Susan P. Douglas, CESA (France)*  
*W. Fred Van Raaij, Univ. of Tilburg (The Netherlands)*

Cypress Room — 9:00 - 10:30

**“CONSUMER SATISFACTION/DIS-SATISFACTION”**

CHAIRMAN: *John E. Swan, Univ. of Alabama, Birmingham*

“Conceptual and Methodological Issues in Consumer Satisfaction Research”

*Stephen A. LaTour, Univ. of Florida*  
*Nancy C. Peat (student), Univ. of Florida*

“Consumer Response to Dissatisfaction With Durable Products”

*Ralph L. Day, Indiana Univ.*  
*Stephen B. Ash, Univ. of Western Ontario*

“Dissatisfaction Attributions and Consumer Complaint Behavior”

*S. Kirshnan (student), Univ. of Pittsburgh*  
*Valerie A. Valle, Univ. of Pittsburgh*

DISCUSSION PAPERS:

*George H. Haines, Univ. of Toronto*  
*E. Jay Russo, Univ. of Chicago*

**(CONTINUES)**

**(CONTINUES)**

**BREAK 10:30 - 10:45**

Banyan East

ATTRIBUTION THEORY MARCHES ON

Live Oak Room

EUROPEAN RESEARCH CONTINUES

**SUNDAY MORNING SESSIONS — II**

Cypress Room — 10:45 - 12:15

**“CHILDREN: FROM TOTS TO TARTS”**

CHAIRMAN: *Lawrence H. Wortzel, Boston Univ.*

“An Exploratory Study of Toddlers’ Satisfaction With Their Toy Environments”

*Gail Bjorklund, Hood College*  
*Richard Bjorklund, George Mason Univ.*

“Children’s Attitudes Toward Advertising On Television and Radio and In Children’s Magazines and Comic Books”

*Jay D. Lindquist, Western Michigan Univ.*

“The Influence of Children In Family Decision-Making: Parents’ Perceptions”

*Roger L. Jenkins, Univ. of Tennessee*

“Children As Information Sources In The Family Decision To Eat Out”

*James E. Nelson, Montana State Univ.*

DISCUSSION PAPERS:

*John R. Rossiter, Univ. of Pennsylvania*  
*Scott Ward, Harvard Univ.*

**12:15 CONFERENCE ENDS ... HAVE A GREAT YEAR ... SEE YOU IN SAN FRANCISCO IN '79!**



**NOTES**  
9:00 - 10:30

**SUNDAY MORNING SESSIONS — I**

Banyan West — 9:00 - 10:30

**“EMPIRICAL METHODS AND FINDINGS”**

CHAIRMAN: *Luis V. Dominguez*, Case-  
Western Reserve Univ.

“An Empirical Examination of Alternative  
Models For Predicting Consumer Utilization  
of Two Credit Card Systems”

*Elizabeth C. Hirschman*, New York Univ.  
*Rajendra K. Srivastava*, Univ. of Texas  
*Mark I. Alpert*, Univ. of Texas, Austin

“The Role of Sex Role Self-Concept in  
Masculine and Feminine Product  
Perceptions”

*Linda L. Golden*, Univ. of Texas, Austin  
*Neil Allison*, Univ. of Cincinnati  
*Mona Clee* (student), Univ. of Texas,  
Austin

“Understanding and Analyzing Contingency  
Data”

*Edgar A. Pessemier*, Purdue Univ.

DISCUSSION PAPERS:

*Zarrell V. Lambert*, Auburn Univ.  
*William D. Perreault*, Univ. of North  
Carolina

---

**BREAK 10:30 - 10:45**

---

**NOTES**  
10:45 - 12:15

**SUNDAY MORNING SESSIONS — II**

Banyan West — 10:45 - 12:15

**“BROADENING THE CONCEPT OF CON-  
SUMER BEHAVIOR . . . SOME INITIAL  
RESULTS”**

CHAIRMAN: *Michael L. Rothschild*, Univ. of  
Wisconsin

“Explaining Education and Occupational  
Choices: An Exploratory Study”

*Paul N. Bloom*, Univ. of Maryland  
*George Coan* (student), Univ. of  
Maryland

“Organizational Buying Behavior: A Con-  
ceptual View Of The Buying Center As An  
Information Processing Unit”

*James W. Hanson* (student), Univ. of  
Oregon

“Innovative Health Care Alternatives: A  
Model Of The Consumer Behavior Process”

*Stephen J. Miller*, Oklahoma State Univ.  
*William G. Zikmund*, Oklahoma State  
Univ.

“Role of Attribute Generality In Cognitions  
Of Political Candidates”

*Glenn S. Omura*, Michigan State Univ.

“An Exploratory Study Of The Evolution Of  
The Negative Image Of Personal Selling”

*John J. Painter*, Univ. of Utah  
*Richard J. Semenik*, Univ. of Utah



## TABLE OF CONTENTS

Association for Consumer Research Conference and Executive Committees . . . . .	iii
Organizers of Special Sessions and Judges for the Paper Competition . . . . .	iv
Preface . . . . .	v
1978 Conference Program . . . . .	vii
Presidential Address . . . . .	1
Jerome B. Kernan, University of Cincinnati	

### PART I: WORKSHOP AND SPECIAL TOPIC SESSION SUMMARIES

#### A. CURRENT ISSUES

<i>Michael B. Mazis</i>	
Overview of "Can and Should the FTC Restrict Advertising to Children" Workshop . . . . .	3
<i>Scott Ward</i>	
Researcher's Look at the "Kid Vid" Rule: Overview of Session . . . . .	7
<i>R. Bruce Hutton</i>	
Overview of "The Energy Crisis and Consumer Conservation: Current Research and Action Program" Workshop . . . . .	12
<i>J. J. Persensky</i>	
Overview of Workshop "Consumer Research Goes to Washington" . . . . .	15

#### B. FRONTIER TOPICS

<i>Fred Van Raaij</i>	
European Consumer Research: Frontier Issues . . . . .	18
<i>Charles W. King, Douglas J. Tigert and Lawrence J. Ring</i>	
Pragmatic Applications of Consumer Research in Retailing . . . . .	20
<i>E. Thomas Garman</i>	
Overview of the "Home Economics Research: Why Give a Damn?" Workshop . . . . .	27
<i>Fred G. Cutler</i>	
Overview of the "Longitudinal Research Design and Analysis of Panel Data" Workshop . . . . .	31
<i>Rebecca Holman and M. Venkatesan</i>	
Overview of: "Time: The Fundamental Things Apply" . . . . .	34

#### C. SPECIAL TOPIC COMPETITIVE PAPERS

<i>Philip Hendrix, Thomas C. Kinnear and James R. Taylor</i>	
The Allocation of Time by Consumers . . . . .	38
<i>Peter Reingen</i>	
Inducing Compliance With A Request: The List Technique . . . . .	45
<i>Carol Scott and Alice Tybout</i>	
Extending the Self-Perception Explanation: The Effect of Cue Salience on Behavior . . . . .	50

### PART II: COMPETITIVE PAPERS

#### "SEX IN ADVERTISING"

<i>R. Dale Wilson and Noreen K. Moore</i>	
"The Role of Sexually-Oriented Stimuli in Advertising: Theory and Literature Review" . . . . .	55
<i>Donald Sciglimpaglia, Michael A. Belch and Richard F. Cain, Jr.</i>	
"Demographic and Cognitive Factors Influencing Viewers' Evaluations of 'Sexy' Advertisements" . . . . .	62
<i>Mary Lou Roberts and Perri B. Koggan</i>	
"How Should Women be Portrayed in Advertisements? - A Call for Research . . . . .	66

DISCUSSION PAPERS:

*Mark I. Alpert*  
Sex Roles, Sex, and Stereotyping in Advertising: More Questions than Answers . . . . . 73

*Edward C. Strong*  
Prisoner of Sex in Advertising . . . . . 78

"FRONTIER ISSUES IN ADVERTISING"

*Joel Saegert*  
"A Demonstration of Levels-Of-Processing Theory in Memory for Advertisements" . . . . . 82

*Peter H. Webb*  
"Perceptual Discrepancies in the Time Duration and Number of Television Commercials" . . . . . 85

*Edwin C. Hackleman and Subhash C. Jain*  
"An Experimental Analysis of Attitudes Toward Comparison and Non-Comparison Advertising" . . . . . 90

DISCUSSION PAPERS:

*John G. Myers*  
Levels-of-Processing, Perceptual Bias, and Comparison Advertising . . . . . 95

*Alan G. Sawyer*  
Frontiers of Advertising Research . . . . . 99

"PERSONALITY AND CONSUMER TRAITS"

*Claes Fornell and Robert A. Westbrook*  
"An Exploratory Study of Assertiveness, Aggressiveness, and Consumer Complaining Behavior . . . . . 105

*Elizabeth C. Hirschman and Melanie Wallendorf*  
"Correlations Between Three Indicators of Breadth and Variation in Sources of Stimulation" . . . . . 111

*Kenneth A. Coney and Robert R. Harmon*  
"Dogmatism and Innovation: A Situational Perspective" . . . . . 118

DISCUSSION PAPERS:

*Harold H. Kassarjian*  
Personality: The Longest Fad . . . . . 122

*Jerome B. Kernan*  
Personality and Consumer Traits: The Beat Goes On . . . . . 125

"DEMOGRAPHICS, SPENDING, AND LEISURE"

*Johan Arndt*  
"Family Life Cycle As A Determinant of Size and Composition of Household Expenditures . . . . . 128

*E. Laird Landon, Jr., William B. Locander*  
"Family Life Cycle and Leisure Behavior Research" . . . . . 133

*Robert B. Settle, Pamela L. Alreck and Michael A. Belch*  
"Social Class Determinants of Leisure Activity" . . . . . 139

DISCUSSION PAPERS:

*Robert Ferber*  
Comments on Papers on Life Cycle Analysis . . . . . 146

*Charles W. King*  
Demographics, Spending and Leisure: A Critique . . . . . 149

"BACK TO THE BASICS"

*K. W. Kendall and Ian Fenwick*  
"What Do You Learn Standing in a Supermarket Aisle? . . . . . 153

*G. A. Mauser*  
"Allison & Uhl Revisited: The Effects of Taste and Brand Name on Perceptions and Preferences" . . . . . 161

<i>David C. Arch</i>		
"Pupil Dilation Measures in Consumer Research: Applications and Limitations" . . . . .		166
DISCUSSION PAPERS:		
<i>James F. Engel</i>		
Back to the Basics: Discussion . . . . .		169
<i>Joseph W. Newman</i>		
Discussion of "Back to the Basics" Research Papers . . . . .		171
"INVOLVEMENT"		
<i>John L. Lastovicka</i>		
"Questioning the Concept of Involvement Defined Product Classes" . . . . .		174
<i>Larry M. Newman and Ira J. Dolich</i>		
"An Examination of Ego-Involvement as a Modifier of Attitude Changes Caused from Product Testing" . . . . .		180
<i>Kjell Grønhaug and Norman Kangun</i>		
"Exploring Generalized and Personalized Beliefs Among Smokers and Non-Smokers: A First Look" . . . . .		184
DISCUSSION PAPERS:		
<i>Andrew A. Mitchell</i>		
Involvement: A Potentially Important Mediator of Consumer Behavior . . . . .		191
<i>Michael L. Ray</i>		
Involvement and Other Variables Mediating Communication Effects as Opposed to Explaining All Consumer Behavior . . . . .		197
"CONSUMER DECISION MAKING"		
<i>P. S. Raju</i>		
"Stimulus-Response Variables in New Product Research" . . . . .		200
<i>John R. Walton and Eric N. Berkowitz</i>		
"The Effects of Choice Complexity and Decision Freedom on Choice Behavior" . . . . .		206
<i>James W. Harvey</i>		
"Evaluative Conflict and Information Search in the Adoption Process" . . . . .		209
DISCUSSION PAPERS:		
<i>James R. Bettman</i>		
Issues in Research on Consumer Choice . . . . .		214
<i>Peter L. Wright</i>		
Decision Process Theory and Research . . . . .		218
"INFORMATION PROCESSING AND EVOKED SETS"		
<i>Frederick E. May</i>		
"Evoked Set Formation and Composition: The Learning and Information Processing Hypotheses" . . . . .		222
<i>Thomas L. Parkinson and Michael Reilly</i>		
"An Information Processing Approach to Evoked Set Formation" . . . . .		227
<i>Joseph J. Belonax, Jr.</i>		
"Decision Rule Uncertainty, Evoked Set Size and Task Difficulty as a Function of Number Choice Criteria and Information Variability" . . . . .		232
DISCUSSION PAPER:		
<i>James H. Myers</i>		
Methodological Issues in Evoked Set Formation and Composition . . . . .		236
"FURTHER DEBATE ON MULTIATTRIBUTE ATTITUDES"		
<i>E. H. Bonfield</i>		
"A Comment on the State of Attitude Measurement in Consumer Research: A Polemic" . . . . .		238
<i>Joel Huber and Robert P. Leone</i>		
"A Note on the Fishbein Attitude Model as an Expected Utility Model" . . . . .		245

<i>Arun K. Jain, Vijay Mahajan and Naresh K. Malhotra</i> "Multiattribute Preference Models for Consumer Research: A Synthesis" . . . . .	248
DISCUSSION PAPERS:	
<i>Olli T. Ahtola</i> Some Issues in Attitude Theory and Measurement . . . . .	253
<i>Sadaomi Oshikawa</i> Attitude Models in Consumer Research . . . . .	256
"ATTRIBUTES: IMPORTANT, SALIENT AND IDEAL"	
<i>James H. Myers and Richard F. Chay</i> "A Comparison of Two Methods For Determining Optimum Levels of Product Characteristics" . . . . .	259
<i>John A. Quelch</i> "Behavioral Measurement of the Relative Importance of Attribute-Related Information Cues: The Case of Cold Breakfast Cereals" . . . . .	263
<i>Jerry C. Olson and Aydin Muderrisoglu</i> "The Stability of Responses Obtained By Free Elicitations: Implications for Measuring Attribute Salience and Memory Structure" . . . . .	269
DISCUSSION PAPERS:	
<i>Richard P. Bagozzi</i> Discussion Comments: "Attributes: Important, Salient, and Ideal" . . . . .	276
<i>Donald R. Lehmann</i> Searching for Important Attributes and Appropriate Levels . . . . .	280
"KEY QUESTIONS IN VALIDITY"	
<i>Philippe Cattin</i> "On the Use of Formulas of the Predictive Validity of Regression in Consumer Research" . . . . .	284
<i>George John and Torger Reve</i> "Construct Validation in Marketing: A Comparison of Methods in Assessing the Validity of the Affective, Conative, and Cognitive Components of Attitudes" . . . . .	288
<i>Richard P. Bagozzi and Robert E. Burnkrant</i> "Attitude Measurement and Behavior Change: A Reconsideration of Attitude Organization and Its Relationship to Behavior" . . . . .	295
DISCUSSION PAPERS:	
<i>Joel B. Cohen</i> Exploring Attitude Construct Validity: Or Are We? . . . . .	303
<i>Richard Staelin</i> Structuring Modeling: An Application for Testing Attitude Models and Convergent Validity . . . . .	307
"SITUATIONAL IMPACTS ON CONSUMER PURCHASING: GIFT-GIVING"	
<i>Thomas K. Clarke and Russell W. Belk</i> "The Effects of Product Involvement and Task Definition on Anticipated Consumer Effort . . . . .	313
<i>Sharon K. Banks</i> "Gift-Giving: A Review and an Interactive Paradigm" . . . . .	319
<i>Roger Heeler, June Francis, Chike Okechuku and Stanley Reid</i> "Gift Versus Personal Use Brand Selection" . . . . .	325
DISCUSSION PAPERS:	
<i>Richard J. Lutz</i> Consumer Gift-Giving: Opening the Black Box . . . . .	329
<i>Douglas J. Tigert</i> Three Papers on Gift-Giving: A Comment . . . . .	332

"PSYCHOLOGICAL CONTRIBUTIONS"

*Jonathan Gutman and Donald E. Vinson*  
 "Value Structures and Consumer Behavior" . . . . . 335

*John L. Swasy*  
 "Measuring the Bases of Social Power" . . . . . 340

*Michael K. Mills and Thomas V. Bonoma*  
 "Deviant Consumer Behavior: A Different View" . . . . . 347

DISCUSSION PAPERS:

*Stewart Bither*  
 A Comment on Value Structures, Power-Content Interaction Theory and the Measurement of Social Power. . . . . 353

*Jagdish N. Sheth*  
 Discussion of Psychological Contributions . . . . . 355

"FAMILY DYNAMICS"

*George P. Moschis and Roy L. Moore*  
 "Family Communication and Consumer Socialization" . . . . . 359

*Susan P. Douglas*  
 "A Cross-National Exploration of Husband-Wife Involvement in Selected Household Activities" . . . . . 364

*Alvin C. Burns and David J. Ortinau*  
 "Underlying Perceptual Patterns in Husband and Wife Purchase Decision Influence Assessments" . . . . . 372

DISCUSSION PAPER:

*Donald H. Granbois*  
 Wives, Husbands, and Children: Three Studies of Consumer Roles . . . . . 377

"SOCIOLOGICAL PERSPECTIVES"

*Margaret DeBell and Rachel Dardis*  
 "Extending Product Life: Technology Isn't the Only Issue" . . . . . 381

*William O. Bearden, A. William Gustafson and J. Barry Mason*  
 "A Path-Analytic Investigation of Life Satisfaction Among Elderly Consumers" . . . . . 386

*Theodore F. Smith, Richard W. Olshavsky and Michael F. Smith*  
 "Smoking Behaviors as a Diffusion Process Within Age Cohort Groups: An Application of the Societal Marketing Concept: . . . . . 392

DISCUSSION PAPERS:

*Thomas S. Robertson*  
 Discussion Paper: Sociological Perspectives . . . . . 396

*Gerald Zaltman*  
 Discussion Paper: Sociological Perspectives . . . . . 398

"CHILDREN: FROM TOTS TO TARTS"

*Gail Bjorklund and Richard Bjorklund*  
 "An Exploratory Study of Toddlers' Satisfaction With Their Toy Environments" . . . . . 400

*Jay D. Lindquist*  
 "Children's Attitudes Toward Advertising On Television and Radio and in Children's Magazines and Comic Books" . . . . . 407

*Roger L. Jenkins*  
 "The Influence of Children In Family Decision-Making: Parents' Perceptions" . . . . . 413

*James E. Nelson*  
 "Children As Information Sources In The Family Decision To Eat Out" . . . . . 419

DISCUSSION PAPERS:

*John R. Rossiter*  
 Children's Consumer Research: A Call for Rigor . . . . . 424

<i>Scott Ward</i> Research on Marketing and Children: Upside or Downside on the Product Life Cycle? . . . . .	427
<b>"CONSUMER SATISFACTION/DISSATISFACTION"</b>	
<i>Stephen A. LaTour and Nancy C. Peat</i> "Conceptual and Methodological Issues in Consumer Satisfaction Research" . . . . .	431
<i>Ralph L. Day and Stephen B. Ash</i> "Consumer Response to Dissatisfaction With Durable Products" . . . . .	438
<i>S. Kirshnan and Valerie A. Valle</i> "Dissatisfaction Attributions and Consumer Complaint Behavior" . . . . .	445
<b>DISCUSSION PAPERS:</b>	
<i>George H. Haines</i> Three Papers on Consumer Satisfaction/Dissatisfaction: A Comment . . . . .	450
<i>E. Jay Russo</i> Consumer Satisfaction/Dissatisfaction: An Outsider's View . . . . .	453
<b>"CONSUMER POLICY PROPOSALS"</b>	
<i>George P. Moschis</i> "Formal Consumer Education: An Empirical Assessment" . . . . .	456
<i>Stephen A. Goodwin, Vijay Mahajan and Bhal J. Bhatt</i> "On Consumer Dissatisfaction: Consumer Arbitration as an Alternative Dispute Resolution Mechanism . . . . .	460
<i>J. R. Brent Ritchie and John D. Claxton</i> "In Search of Actions to Reduce Consumer Shopping Problems" . . . . .	466
<b>DISCUSSION PAPERS:</b>	
<i>E. Scott Maynes</i> Comments on Papers on Consumer Policy Proposals . . . . .	472
<i>Hans B. Thorelli</i> "Consumer Policy Proposals" Discussion Paper . . . . .	477
<b>"THE CONSUMER ENVIRONMENT: REGULATION AND RESEARCH"</b>	
<i>Steven E. Permut</i> "Research on Consumer Information: Public Sector Perspectives" . . . . .	481
<i>John L. Evans, C. Dennis Anderson and L. G. McCabe</i> "Consumer Research Inputs Into Public Policy Decision Making: The Role of Canada's Consumer Research and Evaluation Branch" . . . . .	488
<i>Marc A. Grainer, Kathleen A. McEvoy and Donald W. King</i> "Consumer Problems and Complaints: A National View" . . . . .	494
<b>DISCUSSION PAPERS:</b>	
<i>Stephen A. Greyser</i> The Consumer Environment: Regulation and Research - Discussion . . . . .	501
<i>Ivan L. Preston</i> The Consumer Environment: Regulation and Research: Discussion of Three Papers . . . . .	505
<b>"RESEARCH CONTRIBUTIONS TO PUBLIC POLICY"</b>	
<i>Philip C. Burger and Alladi Venkatesh</i> "A Study of Public Policy Impact on Consumer Decision-Making" . . . . .	508
<i>Jacques C. Bourgeois</i> "A Measure for Market Delineation" . . . . .	514
<i>Linda L. Golden, John F. Betak and Mark I. Alpert</i> "Attracting Potential Switchers to Mass Transit: Mode Choice as a Multiattribute Decision Model" . . . . .	519



DISCUSSION PAPERS:

*Ivan Ross*  
Research Contributions to Public Policy" Discussion Comments . . . . . 526

*Frederick D. Sturdivant*  
Marketing and Public Policy Contributions: Comments on Three Papers . . . . . 529

"ISSUES IN SURVEY MEASUREMENT"

*Abraham Horowitz and Thomas F. Golob*  
"Survey Data Reliability Effects on Results of Consumer Preference Analyses" . . . . . 532

*Roger Best, Del I. Hawkins and Gerald Albaum*  
"The Effect of Varying Response Intervals on the Stability of Factor Solutions of Rating Scale Data. . . . . 539

*Kenneth E. Miller and Marilyn Turner*  
"The Effects of Sponsorship on Mail Survey Response and Evaluation Bias" . . . . . 542

DISCUSSION PAPERS:

*Gordon G. Bechtel*  
Discussion Paper: Issues in Survey Measurement . . . . . 545

*Paul E. Green*  
Discussion Paper: Issues in Survey Measurement . . . . . 548

"SIGNIFICANT PROBLEMS IN ESTIMATION"

*James G. Barnes and Jacques C. Bourgeois*  
"Estimating the Effects of Advertising: Application To A Social Problem" . . . . . 550

*Richard C. Reizenstein and David J. Barnaby*  
"Assessing the Potential Effects of Differential Price Increases on Gasoline Usage" . . . . . 557

*F. Robert Dwyer*  
"Drug Compliance And The Neglected Concern For Validity" . . . . . 564

DISCUSSION PAPER:

*Albert R. Wildt*  
Statistical Estimation and Consumer Research . . . . . 569

"METHODOLOGICAL DEVELOPMENTS"

*James McCullough, Charlene S. Martinsen and Linda Sceurman*  
"Changes in Consumer Perceptions: The Impact of Testing Conditions on Perceptions of Branded Products" . . . . . 574

*Joel Huber and Murphy A. Sewall*  
"Covariance Bias of Thurstone Case V Scaling as Applied to Consumer Preferences and Purchase Intentions" . . . . . 578

*James B. Wiley and Robert Bushnell*  
"Market Shares Estimates Based on Conjoint Analysis of Concepts" . . . . . 582

DISCUSSION PAPERS:

*Robert Blattberg*  
Discussant's Comments . . . . . 587

*James L. Ginter*  
A Discussion of "Methodological Developments" . . . . . 589

"EMPIRICAL METHODS AND FINDINGS"

*Elizabeth Hirschman, Rajendra K. Srivastava and Mark I. Alpert*  
"An Empirical Examination of Alternative Models For Predicting Consumer Utilization of Two Credit Card Systems" . . . . . 592

*Linda L. Golden, Neil Allison and Mona Clee*  
"The Role of Sex Role Self-Concept in Masculine and Feminine Product Perceptions" . . . . . 599

<i>Edgar A. Pessemier</i>	
"Understanding and Analyzing Contingency Data" . . . . .	606
<b>DISCUSSION PAPERS</b>	
<i>Zarrell V. Lambert</i>	
Some Observations and Questions That Might Be Helpful In Future Research . . . . .	611
<i>William D. Perreault</i>	
Issues Concerning Empirical Methods and Findings in Consumer Research . . . . .	614
 <b>"BROADENING THE CONCEPT OF CONSUMER BEHAVIOR...SOME INITIAL RESULTS"</b>	
<i>Paul N. Bloom and George Coan</i>	
"Explaining Education and Occupational Choices: An Exploratory Study" . . . . .	617
<i>James W. Hanson</i>	
"Organizational Buying Behavior: A Conceptual View Of The Buying Center As An Information Processing Unit" . . . . .	622
<i>Stephen J. Miller and William G. Zikmund</i>	
"Innovative Health Care Alternatives: A Model Of The Consumer Behavior Process" . . . . .	628
<i>Glenn S. Omura</i>	
"Role of Attribute Generality In Cognitions Of Political Candidates" . . . . .	635
<i>John J. Painter and Richard J. Semenik</i>	
"An Exploratory Study Of The Evolution Of The Negative Image Of Personal Selling . . . . .	641
 Author Index . . . . .	 645
Order Form . . . . .	647

PRESIDENTIAL ADDRESS:  
CONSUMER RESEARCH AND THE PUBLIC PURPOSE

Jerome B. Kernan, University of Cincinnati

In an effort to march to the cadence established by my immediate predecessor, I shall foist no pedantic harangue on you. Rather, permit me a few comments--opionated, to be sure--on what we as consumer researchers do and to compare that with what we might do to better serve the public purpose.

First, what is the public purpose? While we might disagree on emphasis, I suspect most of us could accept consumers as the focal point of this question and the maintenance of their "rights" as reasonable criteria for consumer research's contribution. In other words, whatever is the ultimate public purpose, some part of it having to do with consumers is measured by the extent to which their legitimate expectations are satisfied by the research establishment. Such "rights" were enunciated by JFK in a 1962 message to Congress as the rights to safety, to be informed, to choose, and to be heard. It requires no ideological commitment to allow that these are not putative rights. Indeed, all members of this organization--whether we work for corporations, government agencies or universities--have a stake in the consumer's welfare. Our jobs depend on it.

The Is vs. the Might Be

Allow me now to hop, skip and jump over the mentalities that seem to pervade our respective efforts. If we work for a corporation (unless we're in an extremely exceptional ambience) our research is problem solving--how to make things work better. If we work for a government agency our research tends toward how to support a policymaker's position. And if we work for a university our efforts are often--as not guided by the preferences of editorial boards. Let me rush to assert that there is nothing whatever wrong with any of this. Corporations need their problems solved; policymakers need research support and academic journals do a generally excellent job of culling the uninspired effort. What bothers me is that with no appreciably greater effort we could reap considerably greater benefits. But it requires a bit of collusion.

Before I am accused of advocating felonious behavior, let me elaborate. Corporate researchers seek to predict behavior; government researchers want to prescribe it; and university researchers try to explain it. These may be turgid delineations, but they're not untrue. In any event, it would seem that everyone should wish to understand consumer behavior. And to do that (unless one simply repudiates the scientific method) it is essential that we first describe what consumers do. On that foundation, it is possible to lay an integral framework of explanation, prediction and prescription or, to put it another way, to orchestrate the natural talents of academia, government and industry so as to enhance consumer welfare.

The Opportunity

We're awfully parochial. The cultures indigenous to the academic, corporate, and government researcher not only prescribe what each should do but also enforce these prescriptions with reward/punishment systems. Thus, the academic pursues promotion and tenure with theoretic contributions, the corporate researcher chases a vice presidency and its perquisites on a profit rocket,

and those in government seek to transcend elected administrations with ineluctable research capabilities. We're comfortably settled into our respective systems; the inculcation is insidious.

But we can prosper in our respective cultures and cooperate. We need yield no integrity in order to systematize our efforts. Indeed, the only thing that we need change is our attitude, our approach to research. And the result, simply by putting the existing pieces together in a more efficacious way, can be a greater payoff to all concerned--including consumers. It's a deal "you can't refuse."

For Instance

What I propose is nothing novel; to my knowledge it was advocated at least some twenty years ago by Coombs, Raiffa and Thrall (1954). For many reasons, however, it is rarely practiced--especially by academic researchers. If you doubt that, read a sample of journal articles; you'll be convinced.

Assuming that we've all seen the light as a result of Jacoby's (1976) comments, a typical piece of academic research starts with a theory (remember that commandment from graduate school?)--and that's where the source of the trouble resides. Never mind whether it's a high assay theory, so long as the researcher is intrigued by it. This theory, under the best of circumstances, is then "tested" in an experimental setting so antiseptically sterile that internal validity can be scooped up off the floor. Some ANOVA model is then rolled in and--surprise--statistically significant results are found. If you do an especially good job you get three-asterisk results ( $p=.001$ ), which seems not terribly unlike the stars grammar-school teachers paste on their pupils' papers for work well done.

These highly significant results (main effects, interactions, whatever) are then "explained" in a highly-tortured "discussion" section of a journal article which, in the typical case, reduces to a thinly-veiled attempt by the author to coopt the reader into the fantasy that prompted the research in the first place. That such research has no pragmatic value is not cause for scorn; it's either exploratory ("thanks for letting me ignore all those assumptions") or theoretical ("let's pretend").

Let me reiterate my earlier comment. Academic research is not bad. But it can be made so much better if approached in a more systemic fashion. And rather than attempt a didactic epistle on the nature of "better," let me offer you an example. Cialdini et al (1978) have approached research in this fashion (they call it "full-cycle" research) and thereby demonstrated how viable systemic research can be.

They begin with the everyday practice of low-balling--for example, where a car salesperson offers a prospective buyer an engagingly low price in order to induce a commitment to purchase and then removes that incentive ("Sorry, but the sales manager won't approve this price"; etc). Now that's a practice to which we can all relate. It's a business tactic (you decide the valence), a source of consumer frustration (hello, government), and an interesting manifestation of psychologi-

cal compliance (why do consumers fall for this ploy?).

Space prohibits a definitive account of the Cialdini research, but the highlights are that they demonstrated that low-balling: (1) in fact works (it isn't a delusion of car dealers); (2) is conceptually different from the foot-in-the-door technique (which escalates the target task, rather than its cost); and (3) is effective only when prospects perceive a high degree of volition in their initial decision ambience. Three potential explanations were refuted experimentally--behavioral engulfing theory (Heider, 1958), dissonance theory (Brehm, 1956), and self-perception theory (Bem, 1972)--with only a commitment formulation (Kiesler, 1971) satisfactorily accounting for the pattern of experiental results. In simple terms, what this means is that a person who makes a free choice feels a responsibility to perform whatever action is attendant to that choice, even when the cost of that action increases beyond initial expectations.

Unpublished data support these findings. But more importantly, consider what this research represents. Never mind low-balling, per se; it is simply one of many compliance-inducing techniques experienced by consumers in day-to-day situations. The systemic, the "full-cycle" notion to be emphasized here is that the research begins with--is inspired by--an ordinary, everyday phenomenon. A theory to account for that phenomenon is then invoked and, appropriately, inspires the experimental design. The results thus lose no internal validity, but gain at least a foothold on external, or "ecological" validity. People begin to care about statistical significance because it applies to relationships that are neither trivial nor sterile.

So academic research (as demonstrated by academic research) can serve a purpose other than "look how smart I am." But university researchers need help.

#### Impediments

Corporate and government researchers frequently ask academics whether there is a "theory" to account for "this or that." A wholly legitimate inquiry--indeed, one that is needed in abundance! An alarming number of academics seem either ignorant of or oblivious to the nature of "this or that," however. If they are to complement their corporate and governmental counterparts' efforts, then, if they are to do "relevant" research, they need more insights into the texture of "this or that." Tell the academic, for example, when his/her experimental design is so tight that it's trivial. We're past the stage where we can afford demonstration exercises. We should be helping, not ignoring one another.

One of the principal reasons for ACR's existence is to prevent our different research cultures from hardening into a cold tyranny. We've been moderately successful. I'm no Pollyanna in the sense that I believe that we're all likely to come together tomorrow. Too much argues against that. But isn't it curious how the superior people seem to manage in spite of the obstacles? Marketing folk--principally through MSI--appear to have set up both a structure and incentive to move in this direction (Greyser, 1978). Can we afford to do less?

#### References

D. J. Bem, "Self Perception Theory," in Advances in Experimental Social Psychology, Vol. 6, Leonard Berkowitz, ed. (New York: Academic Press, Inc., 1972), 1-62.

Jack Brehm, "Postdecision Changes in the Desirability of Alternatives," Journal of Abnormal and Social Psychology, 52 (May 1956), 384-389.

Robert B. Cialdini, John T. Cacioppo, Rodney Bassett and John A. Miller, "Low-Ball Procedure for Producing Compliance: Commitment then Cost," Journal of Personality and Social Psychology, 36 (May 1978), 463-476.

Clyde H. Coombs, Howard Raiffa and Robert M. Thrall, "Some Views on Mathematical Models and Measurement Theory," in Decision Processes, Robert M. Thrall, Clyde H. Coombs and Robert L. Davis, eds. (New York: John Wiley & Sons, Inc., 1954), 19-37.

Stephen A. Greyser, "Academic Research Marketing Managers Can Use," Journal of Advertising Research, 18 (April 1978), 9-14.

Jacob Jacoby, "Presidential Address: Consumer Research, Telling It Like It Is," in Advances in Consumer Research, Vol. 3, Beverlee B. Anderson, ed. (Cincinnati: Association for Consumer Research, 1976), 1-11.

Charles A. Kiesler, The Psychology of Commitment. (New York: Academic Press, Inc., 1971).

OVERVIEW OF "CAN AND SHOULD THE FTC RESTRICT  
ADVERTISING TO CHILDREN" WORKSHOP

Michael B. Mazis, Federal Trade Commission

Introduction

This session was organized to present to the consumer research community the views of leading figures in the debate over children's advertising regulation. The Federal Trade Commission's children's advertising proceeding is particularly significant to consumer researchers because after a number of years of limited research utilization at FTC, behavioral science research now occupies a central role in the policy debate. Given the importance of research findings in the proceeding, eight key individuals, who have been associated with the children's advertising controversy over the past decade, agreed to address an ACR audience. The participants shared the common belief that the research community should be apprised of the issues underlying the trade regulation rulemaking proceedings, which were scheduled to commence hearings (and did) in January, 1979.

The Commission's determination to undertake this rule-making proceeding arises from consideration of two petitions received in April, 1977 from Action for Children's Television and the Center for Science in the Public Interest. While slightly different in approach, these two petitions request promulgation of a trade regulation rule prohibiting the television advertising of sugared products to children.

Following receipt of these petitions, Bureau of Consumer Protection staff undertook an inquiry into the factual and legal issues raised by the petitions. These efforts culminated in a document entitled "Staff Report on Television Advertising to Children." The two petitions and the Report suggest that televised advertising of any product directed to young children who are too young to understand the selling purpose of, or otherwise comprehend or evaluate, commercials (i.e., under the age of eight) may be unfair and deceptive within the meaning of Section 5 of the Federal Trade Commission Act, requiring appropriate remedy. The Report also suggests that current televised advertising of sugared products directed to older children (i.e., between ages of 8 and 12) may be unfair and deceptive, again requiring appropriate remedy.

The staff's recommended approach consisted of a "package" of remedies which included the following three elements:

1. Ban all televised advertising for any product which is directed to, or seen by, audiences composed of a significant proportion of children who are too young to understand the selling purpose of or otherwise comprehend or evaluate the advertising;
2. Ban televised advertising for sugared food products directed to, or seen by, audiences composed of a significant proportion of older children, the consumption of which products poses the most serious dental health risks;
3. Require televised advertising for sugared food products not included in Paragraph (2), which is directed to, or seen by, audiences composed of a significant proportion of older children, to be balanced by nutritional and/or health disclosures funded by advertisers.

In response to the staff's recommendation, the Commission invited comment in an April 28, 1978, Federal Reg-

ister announcement on the advisability and manner of implementation of a rule which would include the preceding three elements. In addition, the Commission sought comment on the appropriateness and workability of alternative remedial approaches not contained in the staff report.

A distinguished panel of experts, representing the full spectrum of views on these proposals, was assembled to state their positions and to interact with the consumer research audience. In order of their presentations, the speakers were:

- Tracy A. Westen, Federal Trade Commission
- Robert B. Choate, Council on Children, Media and Merchandising
- John A. Dimling, National Association of Broadcasters
- Seymour Banks, Leo Burnett Advertising
- Gerald J. Thain, University of Wisconsin Law School
- Stanley Cohen, Advertising Age
- Fletcher C. Waller, General Mills, Inc.
- William Van Brunt, Hershey Foods Corp.

The panel was organized by Michael B. Mazis, Office of Policy Planning and Evaluation, Federal Trade Commission, who served as chair for the session. Dr. Mazis introduced the participants, whose remarks appear in summary form below.

Tracy A. Westen

Mr. Westen has been deputy director of the Federal Trade Commission's Bureau of Consumer Protection since August 1977, and has been one of the architects of the proposed children's advertising rule. His remarks were designed to set forth the facts underlying the FTC's children's advertising proposal and to place the proposal in a legal context.

Westen cited a number of "facts" which led staff to consider some action to remedy the children's advertising "problem." Children between the ages of two and eleven see enormous amounts of advertising--about 25 hours per week is spent watching television and approximately 20,000 ads are seen by the average child in a year. In addition, there is considerable research evidence that young children: can't distinguish between programs and commercials; do not understand the selling intent of commercials; cannot distinguish between fantasy and reality.

Mr. Westen also discussed the problems associated with the advertising of sugared products to children. Children see about 7,000 advertisements for highly sugared products each year, but children between ages seven and twelve have difficulty balancing appeals for the immediate gratification afforded by eating highly sugared products with the long-term dental risks. By age two, about one-half of children have diseased gums and decayed teeth.

The unique legal treatment of children was discussed also. Westen suggested that the Supreme Court has ruled that even fundamental civil rights may be withheld from young children who lack the maturity to exercise them. The ability to run for public office and to

cast a ballot for candidates of one's choice, for example, are vital Constitutional rights. Yet the Constitution itself withholds these rights from individuals below a certain age. And in the commercial marketplace, the courts have voided contracts made with children and restrained the ability of minors to purchase guns, liquor, and cigarettes--products which are felt to be dangerous in immature hands.

Thus, the notion that children often need special protection from the consequences of their inexperience and immaturity is sanctioned by the Constitution and the decisions of the Supreme Court. The thrust of the FTC's proposals is to withhold advertising from pre-schoolers and other children who do not understand what advertising is, and provide older children with both sides of important nutritional controversies. This approach, Mr. Westen stated, appears consistent with decades of Constitutional precedent.

In the case of children's advertising, Mr. Westen emphasized, there is reason to believe that advertisers are engaging in deception through omission of a material fact. Deception occurs because of the immaturity of the audience (i.e., children may be unable to "fill in the blanks" in some ads) and the complexity of the product (i.e., children may not understand the consequences of using a potentially dangerous product). While there is sufficient evidence to open the inquiry, Westen stressed that the Commission is still actively considering a wide range of possible options.

Robert B. Choate

Mr. Choate is the President of the Council on Children, Media and Merchandising--an organization dedicated to studying advertising practices to children and to educating children and their parents about present marketing practices so that they may become more prudent consumers.

The opening portion of Mr. Choate's remarks were devoted to a discussion of Congressional lobbying efforts over proposed children's advertising regulations. He expressed great concern about "overt corporate lobbying against supposed independent agencies." In Choate's view, "corporate lobbyists are trying to emasculate the FTC in its actions vis-a-vis children and advertising."

The proposed children's advertising regulation, Choate asserts, is of far greater significance than the stated issues. Previously, there was insufficient attention to the unequal bargaining power between advertisers and imprudent parties. The proceeding involves a possible alteration of commercial law principles. The real issue is: Should commercial law, dealing with transactions between prudent advertisers and "reasonably imprudent" children, be updated to protect children from television sales pitches?

Mr. Choate also elaborated on Mr. Westen's discussion of television advertising content. He stated that there is an imbalance in the products advertised to children on commercial television. The private enterprise system emphasizes a limited "menu" of highly sugared foods and drinks as opposed to fresh fruits and vegetables or other nutritious foods. The FTC's inquiry is designed to cope with this advertising "imbalance." For example, public service messages to children or to both children and parents are one important method of partially shifting the advertising "imbalance" in the direction of encouraging the purchase and consumption of nutritious foods.

John A. Dimling, Jr.

John A. Dimling, Jr. is Vice-President and Director of Research for the National Association of Broadcasters (NAB). Dimling briefly summarized his perception of the logic of FTC's children's advertising proposals. He stated that the proposed rule rests on the following arguments: (1) per capita sugar consumption has been increasing; (2) sugar (sucrose) causes tooth decay; (3) television advertising to children emphasizes that "sweetness is good" thereby increasing sugar consumption; (4) as a result, we should restrict the number of advertisements directed toward children.

However, the consumption of sucrose, the sugar in candy and pre-sweetened cereals, has remained rather constant. In addition, even if children switched from pre-sweetened cereals to cereals with small amounts of added sugar, they tend to add just as much sugar as exists in pre-sweetened cereals.

In addition, there is very little evidence that eating the most heavily advertised products will cause tooth decay. Pre-sweetened cereal consumed with milk has not been found to have an impact on tooth decay. Also, there is no evidence of a relationship between heavy exposure to television commercials and the incidence of tooth decay. There is some evidence, however, that children are aware that fresh fruits and vegetables are more nutritious than candy bars and other highly sugared foods; as a result the purpose and goal of any FTC action does not appear to be based on a logical foundation.

Finally, Dimling emphasized that the proposed FTC regulations are likely to be counterproductive. The research literature suggests that some children are able to understand the selling intent of commercials earlier than eight years of age, which is the age posited in the FTC staff proposal. Although children below some age do not understand selling intent, restricting advertising to this audience will result in denying advertisers the right of free speech to communicate with other audience members. These other members constitute the great majority of the television audience for most children's programs.

As a result of restricting advertising to large numbers of the potential audience, many of the benefits of advertising would be lost. The possible effects of advertising restrictions include: a decline in the quality and quantity of children's programming; a loss of important information communicated to children through advertising; an increase in the price of toys and other children's products; increased costs of marketing new products, raising entry barriers for many firms.

Seymour Banks

Seymour Banks is Vice-President in charge of Media Research, Leo Burnett Advertising Agency. Dr. Banks discussed two topics during his remarks: the validity of FTC's rulemaking efforts and the need for self-regulation of children's advertising.

Banks confidently predicted defeat of the proposed children's advertising TRR because the FTC staff's conclusions are totally "without merit." There are two main facts leading to this conclusion. First, "the FTC does not have the professional competence to serve as a 'national nanny'." In attempting to present a convincing argument for restricting television advertising to children below the age of eight on the grounds that they cannot understand selling intent, staff engaged in a "highly selective use of research material and an out-of-date oversimplified version of Piaget to the exclusion of social learning theories."

Second, there is little evidence that grade school children are unaware of the long-run consequences of the consumption of heavily sugared products. Banks asserted that the FTC will not be able to defend its position strongly enough to justify First Amendment concerns over the protection of advertisers' rights of free speech.

The second portion of Dr. Banks' remarks was devoted to exploring needed improvements in advertising self-regulation. Advertising agencies and media are "willing to adopt new practices when factual evidence suggests that it would be socially beneficial to do so." What is needed at this point is new research to improve children's advertising. Several areas for additional research were offered: (1) Improvement of children's consumer education skills--such as teaching children to distinguish between programs and commercials; (2) Development of methods to maximize discussion of commercials with children and recommendations about use of these procedures to parents, to maximize incidence of this type of discussion; (3) Development of standards which protect children while allowing for appropriate advertising techniques (e.g., fantasy); (4) Inclusion of pro-social messages in advertising.

Gerald J. Thain

Gerald J. Thain is presently Professor of Law, University of Wisconsin Law School. Prior to assuming a faculty position in 1974, Mr. Thain held a number of important posts at the FTC from 1963-74, including attorney-adviser to Commissioner Philip Elman, Director of the National Advertising Regulation Division and Assistant to the Director, Bureau of Consumer Protection. Professor Thain discussed the two major questions facing the panel: Can the FTC restrict television advertising to children? Should the FTC restrict television advertising to children?

In response to the first question, Thain stated that he believes that the FTC can act to restrict commercials oriented to a children's audience. In terms of the FTC's statutory jurisdiction, it has authority over unfair and deceptive trade practices, which could include children's advertising. Therefore, there do not appear to be any legal inhibitions to FTC action. Although the Supreme Court has ruled that commercial speech is now afforded First Amendment protection, there are unlikely to be any First Amendment problems in the FTC's possible advertising restrictions. The Court has consistently ruled that administrative agencies can engage in reasonable restrictions on the time, place and manner of advertising.

Whether the Commission should take any action which would prohibit of severely restrict advertising to children is a more difficult issue with both political and public policy implications. The political question is: Will it be politically possible to alter the system in the United States which now exists and which many think of as given or status quo? The answer will be greatly influenced by citizens' responses and by lobbying efforts of business and other groups.

The policy question is: Will Commission action result in positive benefit to consumers? The answer depends in large measure how we frame the problem. If we accept the status quo as given, we might ask, "What harm is caused by advertising to children on television?" If we are unwilling to accept the premise underlying this question we might ask, "What social good is served by advertising to children on television?" Determining which question is to be asked may determine the outcome of the FTC's inquiry; different questions might lead to different conclusions.

Stanley Cohen

Stanley Cohen is currently Washington Editor of Advertising Age. His weekly column and editorials have expressed both concern over the content of children's advertising and the likely success of FTC's efforts to develop a workable solution.

Mr. Cohen's stirring statement focused on the political environment surrounding FTC's inquiry into children's advertising regulation. He is impressed with the "destructive power" of the issue. The inquiry should not be over regulation of children's advertising, however, but what is good for the child?

Mr. Cohen believes that broadcasters and advertisers are losing sight of the threat to their own medium in an attempt to preserve their access to the child. He predicted continued advertising to children will turn parents against advertisers, and turn kids against television.

Mr. Cohen raised the question: What has this done to our political process? He claimed behind the scenes lobbying is a threat to the political process. The FTC has procedures based on statutes and due process, adopted with the participation of the industry. Yet, industry was unwilling to have regulations considered in the proper way--instead went to Congress, behind the scenes, and talked to politicians, who have no real knowledge of the issues and who reached their conclusions without hearings. He called this "a government of men, not of laws."

What is this process doing to the First Amendment? As a journalist, Mr. Cohen regards the First Amendment as sacred. His friends have been hauled before judges who have commanded them to disclose their sources. He does not want the First Amendment degraded to allow corporations to tell five-year-olds to eat chocolate-covered cornflakes.

The basic issue, from Mr. Cohen's perspective is: How does the broadcaster carry out its responsibility toward children? It's broadcasters who plan programs and accept ads. He suggested that the FTC approach might be "counter-productive" because it diverts attention away from broadcasters' obligation to prepare good programming for children.

Fletcher C. Waller, Jr.

Fletcher C. Waller, Jr., was named Vice-President and Director of Marketing Services for General Mills in August, 1976. In this position, he has responsibilities for a group of service departments that execute marketers programs and consult with marketers on problem solving. Included are The Betty Crocker Food & Nutrition Center, Promotion Services, Advertising Services, Marketing Research, and Marketing Accounting Departments.

Mr. Waller's remarks were addressed primarily to the implicit assumptions underlying the FTC rulemaking efforts. These implicit assumptions as to how we communicate and deal with children in our society must be made explicit to make progress in the policy debate. According to Waller, the FTC assumes that the American family is incapable of healthy interaction between parent and child. He asked whether "irritation of parents" is justification for FTC regulation on what children should and should not see on television. This intra-family conflict should be considered a natural, indeed healthy, part of parenting and should not be used to increase federal regulation.

Waller believes also that the FTC assumes that children under eight years of age do not have adequate nutrition-

al knowledge. However, the concept of "nutritional knowledge" has not been adequately defined.

Finally, the FTC assumes that advertising to children is synonymous with allowing a surrogate salesman into our homes. However, advertising merely communicates the characteristics of products to children so they can express their desires to parents. Children process information in advertising, match their own set of needs, desires, likes and dislikes with the product array, select what they'd like to have and express these interests to parents.

#### William Van Brunt

William Van Brunt is Associate Counsel for Hershey Foods Corporation with primary responsibility for regulatory matters. Mr. Van Brunt address his remarks to the reasons underlying corporate dissatisfaction with the FTC's children's advertising rulemaking proceeding.

The FTC, in view of the First Amendment protection for Free Speech, and the rule of law that the party seeking to change established customs and practices, has the burden of proof of demonstrating the need and justification for such changes, must prove that the advertising of products to children causes a substantial harm.

On a more specific level, the proposed advertising ban for certain sugared products appears completely unjustified. In fact, no substantial body of data exists: to judge the cariogenicity of food products; to indicate that there is a direct relationship between the percentage of sucrose content and cariogenic potential; to suggest that advertising of sugared products causes harm to children in excess of that which results from the sale of these products without advertising support. The rule-making effort exhibits a measure of elitism, which assumes that the government knows what is right for people.

The possible effect of these efforts, according to Van Brunt, is the imposition of additional inflationary costs on consumers. This potential waste of millions of taxpayer dollars is a result of the FTC's apparent failure to consider fully all the relevant facts before proceeding. For example, staff has ignored industry's self-regulation efforts and the possibility of working for improvements through the process of voluntary industry cooperation. By ignoring or distorting the findings of the recent NSF-sponsored report on children in the marketplace, staff has similarly failed to consider important research evidence which is pertinent to potential rulemaking.

#### Discussion

Following the brief presentations of the panelists, a discussion period ensued, with considerable interaction between panelists and the audience until the end of the session. The panelists remained, however, to comment on the next session, which featured prominent researchers discussing children and television advertising.



## RESEARCHERS LOOK AT THE 'KID VID' RULE: OVERVIEW OF SESSION

Scott Ward, Harvard Business School

### Introduction

This panel of leading researchers in the children's area was convened immediately following the initial session on policy issues in children's advertising. Each of our five speakers delivered a brief summary of his or her general position. These were followed by a lively question and response session. The five speakers were:

Charles Atkin, Michigan State University  
Marvin Goldberg, McGill University  
Thomas Robertson, University of Pennsylvania  
Ellen Wartella, Ohio State University  
Scott Ward, Harvard University

Illustrative comments of each speaker are presented in this paper. Due to the complexity of this topic, and its current importance, readers are cautioned to view these as illustrative comments only, and to pursue greater depths in the expanding literature on children's consumer behaviors.

#### Charles K. Atkin

This summary describes the key conclusions drawn from my 70 page statement submitted to the Federal Trade Commission inquiry into children's advertising. The conclusions are based on an extensive review of available research evidence drawn from almost 100 sources, including my own series of investigations conducted over the past six years.

The most important conclusions will be briefly listed. Then, a more detailed statement will be provided for each topic area.

- Children pay attention to about 10,000 commercials per year, including 1,000 Saturday morning ads for toys and food products.
- Children under eight years old display little understanding of the persuasive intent of advertising and tend to trust the claims made in commercials.
- TV advertising plays a dominant role in shaping children's preferences for toy and food products. Children who heavily view TV ads are far more likely to request and to consume such products.
- Children are persuaded to want cereal because a favorite commercial character promotes it or a toy premium is included in the box; nutritional value is not a salient factor.
- TV commercials contribute to intra-family conflict when frequent food and toy requests are rejected by parents. Ads also create some disappointment and irritation for children.
- Parents do not play a strong direct role in educating children about TV advertising, and they support additional regulations and reductions in child-oriented TV advertising.

Among the remedies for problems associated with children's advertising, the evidence indicates several proposals would be effective: a ban on ads directed to young children, restrictions on practices such as fantasy source characters and premium appeals, limitations

on the amount of advertising time, and PSA-style disclosures about nutrition and health.

#### 1. Characteristics of Advertising Directed to Children: Content Analysis Evidence

A growing number of content analysis studies have sampled children's commercials and systematically tabulated the message attributes in terms of type of product, source characteristics, message appeals, information content, and style of presentation.

These studies show that food products, particularly cereals, are most frequently advertised. Toys are heavily promoted in the months leading up to Christmas. The three basic types of sources are fantasy characters, child characters, and adult figures; most performers and announcers are male. The attractiveness of the characters is far more emphasized than their expertise regarding the product.

The message appeals tend to be emotional rather than rational, contain exaggeration of benefits, and emphasize fun themes. Premium offers are featured in a substantial proportion of ads for cereals and eating places.

Information about substantive qualities of the products is seldom promoted in ads. Food ads feature sweetness or chewiness more often than vitamin attributes. Stylistically, ads tend to be fast-paced and light rather than serious in tone.

Thus, the ads don't provide a solid basis for judging the pertinent merits of the product. The product is associated with irrelevant characters, irrelevant premium offers, fast-paced action in a non-serious tone, and implicit emotional claims promising fun and exaggerated satisfaction. The ingredients, cost, and materials are not emphasized nor are disclaimers effectively communicated in many instances.

#### 2. Children's Exposure and Attention to TV Advertising

Children watch slightly more than three hours of commercial television per day, with Saturday morning viewing accounting for less than 10% of their exposure time. Although children can potentially see 20,000 ads per year, attention studies indicate that actual exposure is considerably less than this. Program attention level drops slightly at the onset of an ad, drops somewhat during the course of the commercial, and drops substantially further for subsequent commercials in a block. Thus, during a time slot when children are watching television, they probably pay close attention to about 40% or 50% of the ads that appear during that period. The estimate for the total number of ads closely attended is likely to be less than 10,000 per year; no more than 1,000 Saturday morning commercials are probably viewed.

#### 3. Children's Understanding of the Nature of Advertising

The level of sophistication regarding the nature of advertising increases sharply with age. Children up to age eight display low levels of understanding of selling intent and little conceptual distinction between programming and advertising. As youngsters develop through late childhood, they demonstrate an increasingly ad-

vanced grasp of selling intent, the biased nature of product presentation in an advertising message, and the substantive difference between programs and ads.

#### 4. Belief of Television Commercials

Inferences regarding children's acceptance of advertising claims depends on the measurement procedure employed. When given a dichotomous forced-choice question asking whether ads are true or untrue, the proportion of children who exhibit generalized distrust rises from about one-fourth of the pre-eight year olds to three-fourths of those over ten years old. However, these studies have not tapped intermediate levels of disbelief, and rejection of specific advertisements and persuasive claims isn't prevalent. Although children are skeptical about assertions in commercials for familiar toys, they readily accept technical claims of a medical or nutritional nature. Heavy viewers of commercials are more likely to believe ads than light viewers.

#### 5. Influence of Commercials on Children's Product Preferences

There is ample evidence that TV advertising plays a dominant role in shaping children's product preferences. Both children and mothers cite commercials as the leading source of awareness of preferred toys and foods. Both experiments and surveys show how exposure to advertising increases desire, asking, and consumption of advertised products. Children who heavily view TV advertising are far more likely to request that parents buy food and toy products for them.

Furthermore, the evidence indicates that advertising stimulates higher usage levels for those consumable products available in the home. There is some tendency for this impact to generalize to other brands; thus, generic consumption is stimulated as well as brand preferences.

Among the source and message factors which produce the strongest effects, premium offers and frequency of repetition appear to make some contribution.

#### 6. Learning About Nutrition From Food Advertising

Among the factors that children weigh in selecting cereals, nutritional value does not seem to be a salient dimension. There is some evidence that children are persuaded to want the cereals for non-substantive reasons--because their favorite character likes the cereal or because they will obtain a toy premium in the box--which are the types of appeals featured most prominently in advertising. However, there are indicators that the limited nutrition-relevant information presented in ads can serve to heighten the importance of this attribute of cereals and to make the balanced breakfast more salient--at least for those who pay attention to that brief segment of the commercial.

Some findings suggest that children's beliefs about the nutritional value of various foods and sugar are shaped by advertising. In general, children are not critical evaluators of claims made in food advertising; they tend to accept the validity of strength benefits and attribute competence to cartoon sources.

The balanced breakfast concept, which is mentioned in most cereal ads, is not well understood by younger children. Rather than informing young viewers about the need for other foods at breakfast, this "disclaimer" may actually give the impression that cereal alone is sufficient. In addition, the ads do not seem to provide adequate awareness of the sugar levels in presweetened cereals or the cariogenic risks associated with sugared cereals, compared to other sweet foods.

#### 7. Other Undesirable Consequences of Advertising

The evidence indicates that TV commercials have an effect on intra-family conflict and child unhappiness. First, parents reject approximately one-third to one-half of child requests for products. This leads to parent-child conflict in about half of all families; such arguments are generally mild and infrequent. The evidence shows that children who see the most advertising have considerably more conflict.

About one-third to one-half of the children become unhappy, angry, or disappointed after denials of food and toy requests; again, the rate is considerably higher among those seeing the most advertising. In addition, some children may become dissatisfied when actual products fall short of the advertised image. Advertising interruptions of program enjoyment produce irritation in many viewers.

#### 8. Parental Mediation and Attitudes Regarding Children's Advertising

Parents do not play a strong direct role in educating children about television advertising. Most simply do not watch these ads with their children, and less than half of the mothers say they teach their children about advertising.

Clearly there is strong public support for additional regulations and reductions in child-oriented TV advertising. Recent surveys show that a majority favors an outright ban, representing a shift from studies done earlier in the 1970's. Regulation by the industry or an independent group is more widely favored than government restrictions.

#### 9. Effectiveness of Proposed Restrictions on Advertising

The Federal Trade Commission Staff Report offers five possible remedies for the potential unfairness and deceptiveness of children's advertising. These include (a) disclosures of nutritional or health information within ads, (b) nutrition or health messages outside ads, (c) limitations on the amount of advertising, (d) limitations of advertising techniques, and (e) bans on all ads directed to younger children and on those ads directed to older children for products posing dental health risks.

Affirmative disclosures within advertisements may not be comprehended by younger children; most do not comprehend the standard disclaimers presented in contemporary toy commercials. However, there is limited basis for optimism: tests of modified versions of toy disclaimers, featuring simplified or conspicuous statements, show greater effectiveness; and researchers demonstrated that children are capable of learning nutrition information presented with a graphic device in the form of a stylized robot. Nevertheless, it is doubtful that advertisers would produce disclosures of sufficient length, prominence, and clarity to achieve similar effectiveness. Thus, the disclosure remedy does not appear to be highly promising at this time.

Affirmative disclosures outside advertisements would involve PSA-style messages to counterbalance the generic influence of advertising. There is considerable evidence that current PSA's have an impact on child audiences. Studies show that children pay close attention and learn the content of these messages; behavioral effects have also been demonstrated. Thus, similar types of spot messages promoting nutrition or warning about cavity risks are likely to be an effective remedy.

Limitations on the amount of advertising might involve reductions in advertising time to four or six minutes

per hour on Saturday morning. To calculate the possible impact on stimulation of product desires and undesirable outcomes, the responses of current light vs. heavy viewers can be examined. Children who now watch a light diet of advertising have an exposure rate similar to the proposed reduced level. The research shows that these children, compared to heavily exposed viewers, are less likely to accept commercial claims, request products, consume advertised products, argue with parents over purchases, and experience unhappiness and dissatisfaction due to commercials. Thus, some beneficial outcomes might be expected if time limitations were instituted.

Restrictions on advertising techniques is a more ambiguous issue, since proscribed practices have not been specified. There is evidence that many children believe that fantasy characters in commercials are competent to discuss the merits of food products, and that these characters have an influence on product preferences. In addition, premium offers are effective in selling food products and contribute to parent-child conflict. Techniques such as these might be restricted, along with other practices identified in subsequent investigations. The need for centralized regulations is given greater impetus by the findings that parental teaching about advertising techniques isn't widespread.

Banning of advertising directed to children under eight who are too young to understand and evaluate commercials is the final remedy. Evidence supporting the case for a ban includes findings that children's commercials feature sophisticated techniques designed to appeal to the unique vulnerability of young children, the large quantity of ads viewed by young children, their low level of understanding of selling intent and uncritical acceptance of commercials, the strong impact on young children's product desire and requests to parents, their advertising-based inappropriate reasons for food product preferences, the contribution of advertising to parent-child conflict and child unhappiness, and the minimal role played by parents in mediating advertising.

Marvin E. Goldberg

The comments that follow stem first from my own direct experience in observing 1,000 or so children over a series of six or seven studies. Somehow one senses the importance of this type of first hand knowledge, when even researchers tend to refer to "my son" or "my daughter" in substantiating their judgments regarding TV's impact on children. Second and more substantively, of course, the findings suggested by the program of research Gerry Gorn and I have conducted at McGill leads me to a particular perspective regarding the issues we are discussing here today.

Researchers must be careful to balance the immediate policy questions of the day with broader policy alternatives. The particular administration we deal with today may change tomorrow, and the policy options that might now seem unrealistic, may seem quite plausible to a new administration. Thus, while a researcher interested in focusing on questions of social concern should not divorce himself from the real policy options of the day, he ought not constrain himself totally within the parameters of a given administration. To do so would be too much of a "reactive" position.

The opportunity to be considerably more "proactive" is typically available to the researcher when he simulates alternative experiences for individuals through the manipulation of experimental conditions. While some experimental conditions may reflect specific policy options, the researcher can also creatively structure additional conditions, and in so doing he may actually become a change agent himself. While questions of one's

own values, biases, etc. may arise, I don't think these are critical and can readily be overcome through standard experimental procedures (keeping experimenters blind, etc.). This is more or less the position we took in developing our "Snack Food" study (Goldberg, Gorn, and Gibson, 1978). The conditions we structured included a TV viewing situation in which children saw a neutral animated cartoon program with either commercials for sugared snacks or PSA's for fruits and vegetables, etc. In these conditions we sought to reflect either current reality or one of a number future realities now sought by the FTC. We went still further, however, asking what would be the impact of TV programming that focused on the nutrition issue. Indeed we found this option was most effective in reducing the number of highly sugared foods children selected.

Just as the researcher ought not let government policies totally bound the perspective he uses in framing his research questions, the same caution ought to be a part of his approach to industry. As a case in point, advertisers typically focus on brand-shifting strategies. In their minds advertising to children is largely intended to shift children away from a competitor's candy bar or sugared cereal to the one they are espousing. The researcher must be alert to this limited perspective. The cumulative effect of each advertiser attempting this strategy is an enormous quantity of candy bar and sugared cereal commercials, certainly relative to those available for less sugared or processed foods. If one's perspective relates to the broader question of TV's effects on the child's total diet, the research questions that would follow would likely be quite different from those of the advertising researcher assessing the effectiveness of a given commercial.

It was suggested earlier this morning that the burden for the FTC is essentially to demonstrate that "grade school children are unaware of long run consequences of eating too much sugar". This seems to me to be an overly narrow and cognitive view of how children are influenced by TV commercials. Our research has shown that while it is true that children may know which foods are, or are not, healthy, this is not really the issue. A more fundamental question is whether children's exposure to TV commercials is such that they are ever motivated to use the healthy-unhealthy dimension in considering what foods to eat. Because the vast majority of foods advertised on children's TV are highly sugared, the child's frame of reference is such that after 2-3 hours of viewing per day, the only alternatives that are likely to be evoked are those he has seen on TV, and by and large these are the highly sugared, processed foods. Joan Gussow (1972) makes this point by analogy to adults; she suggests that just as adults tend to reach for a Budweiser or a Seven-up when thirsty and not the tap for water, so too do kids reach for the products they have been exposed to so exclusively on TV. TV precludes children from active consideration of foods other than those they see on TV. Especially with young children, the notion of "monkey see, monkey do" seems a more appropriate description of their behavior than models which posit intervening cognitive considerations of the nutritional value of these foods.

I would hypothesize that once PSA's espousing a variety of less sugared foods were juxtaposed alongside the current array of TV food commercials, children would begin to actively question the relative merits of different kinds of foods. Once provided with a fuller picture of the array of foods available to them, they would become more capable of considering the tradeoffs associated with different food values (nutrition, fun, etc.). Far from leading to some predetermined notion of "what is right for children" this goal of exposing them to a diversity of foods only has balance in the child's diet as an intended outcome. I doubt if many would consider this as an instance of an overly directive or heavy hand

of government.

Thomas S. Robertson

My comments concerning the impact of advertising on children will focus on the question of what constitutes "effects."

The first point to be made is that it is not necessary to show behavioral effects. There may be many reasons why behavior does not occur -- most obviously the fact that parents are generally the purchasers of goods and services. If television advertising changes or enhances children's beliefs or attitudes, then this is indeed an effect. The argument that there is no resulting behavioral manifestation is largely irrelevant as to whether advertising influenced children.

It should not be overlooked that reinforcement is also an effect. The implications of reinforcement are to enhance existing attitudes and habits. Thus, to suggest that there is no "effect" in promoting presweetened cereals, since children already prefer presweetened cereals, is misleading. Such reinforcement may heighten preference for such products or extend this consumption pattern, whereas it might otherwise wear out over time.

It is sometimes suggested that advertising effects are limited, since parents act as mediators. Nevertheless, there is only limited explicit mediation which occurs within the family concerning consumption and there is limited co-viewing of television with children (Adler, et al., 1977). Furthermore, mediation does not always reduce advertising effects in terms of beliefs and attitudes and in at least two studies by Ward and Robertson (1971) and Robertson, Rossiter and Gleason (1979), parental mediation has been shown to build more favorable attitudes toward consumption. Thus, it is inappropriate to argue that advertising influence is unimportant since parents negate such effects.

Finally, it should be noted that almost all research on advertising and children has focused on short-run effects. The more serious issue may well be the long run socialization effects regarding consumption priorities and patterns. Exposure to thousands of commercials per year must have "some kind of effect" on children and change them in ways that we do not really understand. Herein lies a major avenue for future research.

Ellen Wartella

As Gerald Thain mentioned earlier this morning, I think there are two different questions to be addressed in the rulemaking proceedings:

- To what extent does children's exposure to TV advertising lead to harmful effects?
- To what extent does exposure to TV advertising lead to social benefits for children?

Most of the research on the impact of TV advertising on children, it seems to me, has focused on the first question rather than the second. Let's look at the evidence for both questions.

Regarding the first point, consider particularly the cognitive impact of TV advertising. Extant research evidence indicates that children below middle childhood (between ages 7 and 9) have difficulty understanding the persuasive intent of advertising. While there are some problems with the measures used to examine children's understanding of the purpose of advertising, (for example, mainly abstract questions), I think it is doubtful that other measures will diverge from the present

findings. In particular, I would like to point out that the Giannini and Zuckerman study -- which has been referred to earlier as demonstrating that kindergartners do understand the purpose of advertising -- has grave problems. It measures recognition of program and advertisement characters, but recognition of characters is not evidence of understanding persuasive intent. Beyond this, recognition by the children was only at a chance level. In short, I do not find this study convincing or contrary to other research.

Regarding behavioral effects of TV advertising, there is evidence of modeling effects from Atkin's and Goldberg's work. These were briefly highlighted earlier in this session.

Lastly, turning to the second question posed by Thain, it seems to me that few research studies have addressed the issue of what social good is served by advertising to children. Most studies have examined only short-term effects of advertising messages, not long-term effects. In particular, we do not have any studies in the literature which indicate to what degree early exposure to advertising leads to increased consumer skill development later in the child's development. Nor do we know whether early exposure to TV advertising leads to later understanding of the nature of advertising's persuasive intent. Indeed, I do not believe there is even one longitudinal study of TV advertising's impact on children.

Scott Ward

Reacting to the session speakers, and trying to integrate implications of research on advertising and children, I would like to make four points: First, it seems quite clear that younger children are different from older children. That is not very surprising. What is surprising is the argument some appear to be making, namely that younger children "really can" understand advertising, so that no special attention needs to be given to them.

I'll try to make the second point in terms of a question: even if one attempted to hide advertising from the eyes of very young children, does anyone seriously think that primary demand for candies, popcorn, toys, and so forth, would decrease? In this regard, it is important to recognize that reactance to laboratory conditions in much experimental work can lead to shaky external validity of results. This is particularly critical in the sensitive public policy area such as advertising to kids. In one experiment, for example, it was reported that candy advertising increased children's desire for the generic product category (i.e., primary demand) -- rather than impacting on brand-level (selective) demand. This might have been so in the lab, but it flies in the face of primary demand data which show virtually no change in demand in this product category since the turn of the century.

The third point is that some products are not so good for kids. It is not so black and white an issue that one can say that "if it's not good, ban the product; if it's OK to sell it, it's OK to advertise it." Sticky candy, or sticky raisins, both contain sugar, and both stick to the teeth. Given my second point, the question is: what practical, realistic steps can be taken so that kids use products better (e.g., brush their teeth after eating, evaluate products they see in stores, process advertising and packaging better, etc.)?

That leads to my fourth point. Given that banning advertising is not very practical, in that it would only limit brand competition, not impact on primary demand for "bad" products, the most effective route may be to try and train kids to understand products and advertis-

ing better. The question is, can kids be trained, and if so, how? I am less than optimistic about public service, mass media (e.g., PSA) efforts in this direction. There is precious little of the control and precision one needs for effective education. The school curriculum seems to be a much better forum.

In research Dan Wackman has conducted in Minnesota, it appears that even kindergarten age children can be trained to better evaluate advertising. As data in Table 1 show, children exhibited much greater understanding of the concept of commercials, and persuasive intent, after a one-week, specially designed curriculum program. Research several months later shows stability of results, although by this time, the children's progress through developmental stages helps.

Table 1

	Experi- mental Pre-Test	Control	Experi- mental Post-Test
<u>What is a commercial?*</u>			
Persuasive aspect	0%	15%	46%
Information aspect	43	35	30
Other, lower level understanding	38	30	16
Don't know	<u>19</u>	<u>20</u>	<u>8</u>
	100%	100%	100%
n	(20)	(20)	(37)

\* Asked only of those who answered "yes" to the question, "Do you know what a commercial is?"

Why are commercials  
shown on TV?

Make you buy	29%	28%	60%
Other, lower level understanding	27	26	11
Don't know	<u>44</u>	<u>44</u>	<u>29</u>
	100%	100%	100%
n	(41)	(47)	(53)

What do commercials  
want you to do?

Get you to buy	35%	30%	72%
Get you to try	2	4	0
Other, lower level understanding	15	15	6
Don't know	<u>48</u>	<u>51</u>	<u>22</u>
	100%	100%	100%
n	(41)	(47)	(53)

In conclusion, I tentatively offer that education can be effective. There are clearly major problems with a full-scale policy effort along these lines, most notably business' historic cynicism with "consumer education" (pretty well-founded cynicism, in my view) and the FTC's unwillingness to compromise. Nonetheless, I am increasingly becoming personally convinced that this route holds most promise, and that the effort is worth making.

\* \* \* \*

References

- Adler, Richard P., et al., Research on the Effects of Advertising on Children. Washington, D.C.: U. S. Government Printing Office, 1977, Chapter 10.
- Goldberg, Marvin E., Gerald J. Gorn, and Wendy Gibson, "TV Messages for Snack and Breakfast Foods: Do They Influence Children's Preferences?", Journal of Consumer Research, Vol. 5, September 1978, pp. 73-81.
- Gussow, Joan, "Counternutritional Messages of TV Ads Aimed at Children", Journal of Nutrition Education, 4, 1972, pp. 48-52.
- Robertson, Thomas S., John R. Rossiter, and Terry C. Gleason, Televised Medicine Advertising and Children. New York: Praeger, 1979.
- Ward, Scott and Thomas S. Robertson, "Adolescent Attitudes Toward Television Advertising: Preliminary Findings," in Television and Social Behavior, Vol. IV. Washington, D.C.: U.S. Government Printing Office, 1971, pp. 526-553.

OVERVIEW OF "THE ENERGY CRISIS AND CONSUMER CONSERVATION:  
CURRENT RESEARCH AND ACTION PROGRAMS" WORKSHOP

R. Bruce Hutton, Department of Energy/University of Denver

Abstract

This workshop was designed to inform ACR members about the magnitude and breadth of energy problems, to describe pioneering programs in the U.S. and Canada designed to inform and motivate consumers to conserve, and to stimulate interest in energy policy related research.

Introduction

Perspective for the workshop was quickly established by William Wilkie, chairman, who noted that, since the world is facing incipient disaster, it seemed only appropriate that ACR provide a brief session to glance at the Energy Crisis.

Since the Arab oil embargo, efforts in public policy do, in fact, reflect an urgent need to find solutions to energy problems. While early efforts focused on hardware development, it has become increasingly evident that the consumer is equally important in achieving viable solutions to the "energy crisis." Support for this contention is gained when it is realized that 50% of the population do not know the U.S. has to import oil and another 30% do not know how much. These facts coupled with the current beliefs of a substantial part of the consumer population that there really is no energy problem provides an interesting and challenging arena for consumer researchers.

Consequently, the purpose of the workshop was threefold: (1) to inform ACR members about the magnitude of the energy problem, (2) describe pioneering programs from two countries (Canada and U.S.) designed to inform and motivate consumers to conserve energy, and (3) provide background and stimulate interest in consumer research focused on providing input to policymakers facing important and increasingly serious energy problems. To help accomplish these objectives three individuals representing Canadian and U.S. energy departments described current programs and research: Brian Kelly, Canadian Department of Energy, Mines, and Resources; Jeffrey Millstein, U.S. Department of Energy (DOE); and Bruce Hutton, U.S. DOE/University of Denver.

Current Research on Public Attitudes and Behavior

Jeffrey Millstein provided an overview of consumer attitudes and behavior regarding energy conservation and how they have changed over the last five years. This span represents the time since the Arab oil embargo and subsequently when much of the current conservation effort in the U.S. began.

Mechanisms for gathering this descriptive data included 25 national sample surveys and various metropolitan surveys, focus group interviews, and sales data. Purposes for the research were to:

- 1) Monitor consumers' awareness of the need to save;
- 2) Monitor attitudes toward proposed or actual conservation policies, actors in the energy situation, and means to save energy;
- 3) Monitor knowledge of how to save energy and where to get information;
- 4) Monitor energy saving behaviors, both continuing and one time only kinds;

- 5) Help design more effective communications.

In evaluating the results of the research, Millstein reports that energy conservation is basically a two-fold concept involving efficiency and cutbacks. Most Americans appear to be practicing the former. For example, more than 80% of U.S. homes now have some insulation, 50% of the homeowners have added insulation since purchasing their house, and people are buying more efficient cars. Some cutbacks are also being practiced (e.g., 50% cut back lighting, 60% do not cool below 78° in summer). However, many more are not being demonstrated: only 20% of the population carpool, 12% use public transit to and from work, actual highway speeds are increasing, and vacation travel is rising.

Why people have been reluctant to adopt many of these viable alternatives is not fully known. However, research data suggest several possibilities. First, policy proposals least preferred by a majority of consumers are those that hit closest to home. People do not want to pay more in prices or taxes for energy or related comforts, conveniences, and life styles. They certainly do not want to sacrifice them unless they get something of equal or greater value in return.

Second, people do not want to pay higher prices for energy because, for most, the higher prices are perceived to be the energy problem. Consequently, consumers are baffled by proposals to solve the energy problem through the price mechanism (How can you solve high prices by making them even higher?).

Third, most Americans are poorly informed about important dimensions of the energy problem. The lack of information on factors mentioned earlier such as flagging domestic production and dependence on foreign oil has led to a general lack of concern along a number of fronts. For example, only about one-half the population think energy poses a serious problem, 40% are concerned that foreign oil producers will stop shipments again, only 20% are worried about natural gas shortages, and one-third of the people do not like paying higher prices because they believe the whole issue has been fabricated by oil companies and politicians.

Finally, aside from information and education, people need to feel that measures taken to deal with the energy problem should result in equitable sacrifices. This is one important way to help remove doubts and skepticism about the reality of the energy problem.

Canadian Energy Conservation Information Program

Much of the suspected causes for current levels of consumer attitudes, knowledge, and behavior are being addressed in a variety of programs directed by Brian Kelly in Canada. Kelly reported that since February 1975, the Office of Energy Conservation, Department of Energy, Mines, and Resources has operated an active and diverse information program.

Program Review: 1975 - 1978

The introductory phase of the program began February 1975 and ran for one month. Objectives focused on

awareness of the importance of energy efficiencies among all Canadian publics and changing attitudes toward energy use. To accomplish objectives, the plan incorporated four major elements: (1) advertising through newspaper, radio, and television drawing attention to the need for conservation presented in a positive environment; (2) publications designed to transmit "how to" information; (3) promotional activities aimed at specific market segments; (4) public relations programs to precede the advertising campaign utilizing a special "Speaker's Bureau." Overall, this phase focused on the benefits of energy conservation using credible, straight forward, and positive messages. Both a "crisis atmosphere" and "trivial approaches" were avoided.

Phase Two ran for one year, April 1975 to March 1976. The objectives of this phase were a natural outgrowth of the introductory phase including encouragement to adopt less energy consumptive life styles. Goals, however, became more specific (e.g., to effect a reduction in the rate of growth by 2% or a dollar saving of \$400 million). Target groups also became more specific, ranging from single unit households to specific industries to various ethnic groups. The same major elements were used to meet objectives but with some new features. Special emphasis in advertising was given to the rationale for conservation. A slogan was developed: "If you're not part of the solution, you're part of the problem." In addition, advertising focused on a variety of themes including one designed to capitalize on concerns over inflation: "When we save energy, we fight inflation. One solution to two problems." In addition, Fernand Seguin, a well known French popularizer of science, was chosen as spokesperson for French Canada.

Following Phase Two, Phase Three ran from April 1976 to March 1977. Objectives of the program were extended slightly to encompass an understanding of how energy is linked to other problems, to include other agencies and organizations, and to become more action oriented. Three more target groups were added to the list - agriculture, clergy, and citizen's groups. In addition to the same major elements, this phase marked the beginning of a number of "vertical programs" employing a wide range of communication vehicles focused on a specific area. The first such program dealt with re-insulation, and a new slogan evolved relecting a more solution oriented directions: "Energy conservation - be part of the solution." The major emphasis in this phase was on practical steps and self-help (especially through dollar savings). In order to emphasize potential savings, comparison ads were run showing before/after results of re-insulation.

The most recent phase ran between April 1977 and March 1978. Objectives for this phase encompassed those of Phase Three with additional emphasis on implementation. Increased attention was paid to car drivers, industry, children, and the commercial sector as target groups. The same basic elements for meeting objectives was employed with a variety of updates. The addition of special job creation programs and energy conservation was personified in the form of individuals actively working in community programs for the benefit of everyone. This phase also marked the introduction of "The Hottest Show in Earth," a half-hour television program on insulation. This unique film is a mixture of animation, comedy, drama, and popular science designed to appeal to prime time T.V. audiences. The program has proved extremely popular among several audiences.

The plan for Phase Five (1978-79) is to concentrate on four important areas with the same basic elements. One is automobile transportation which represents the most immediate opportunities for conservation. Second is the industrial sector, the largest energy consumer in the economy, accounting for about 30% of Canada's energy consumption. Third is the residential/farm sector which

consumes 20% of the energy budget. Fourth is the commercial sector which consumer only 14% but is the fastest growing sector in terms of energy use.

#### Program Evaluation

Over the course of the Canadian program, Kelly reports various measures have been used for evaluation including wave data, readership studies, coupon responses, demand for publications, press coverage, and opinion polls. As a result, several conclusions were drawn by Kelly summarizing the Candian experience to date:

- (1) The program must be credible, and scare tactics or frivolous approaches do not produce credibility. Straight-forward, factual and interesting approaches have proven most successful.
- (2) Saving money has the greatest appeal both to the public and business. However, a number of rationales should be developed (e.g., energy supply, balance of payments, national security, enhanced life styles, future generations, etc.) in order to provide a broad basis for conservation.
- (3) Present energy conservation as a positive alternative. Energy shortages and suffering do not represent conservation but rather the inevitable result of not conserving.
- (4) Conservation comes in several types or stages, and an incremental approach to addressing them is advised. In the immediate term concern is with reducing outright energy waste and increasing efficiency. Beyond that are a number of more fundamental issues involving value changes, life styles, urban development, etc. leading to what has been termed "The Conservator Society."
- (5) The execution of an information program should be broad but integrated. It should involve media advertising, publications, promotion, public relations, etc. Media vehicles should be mixed. Publications should be in the form of books wherever possible. Cooperative programs are important. Paid advertising is advised. These vehicles, strategies, and activities should evolve over the life of the program from general to increasingly specific messages.

#### Consumer Research and Programs in the DOE

Bruce Hutton noted that the U.S. has a number of programs comparable to Canada's in the area of information dissemination, but they are less well integrated. Hutton concentrated on one area in Conservation where program strategy and objectives are well defined and consumer research is an integral part of both program formulation and evaluation - the Consumer Motivation Branch (CMB).

The overall objective of the CMB is to encourage private sector groups (e.g., financial institutions, retailers, etc.) to voluntarily work with the DOE to test and evaluate approaches which the private sector can later implement to motivate consumers to become most efficient users of energy. While the ultimate target of CMB activities is the consumer, it is the private sector which is the focus of CMB actions and the mechanism through which objectives will be achieved.

#### Feedback: A Motivation and Information Program

One example where the above strategy has been implemented is the area of feedback. A great deal of research has shown that giving people immediate and understandable feedback on the effects of their actions enables them to better control those actions such as a dieter is assisted by a scale. In research funded by the CMB and conducted by Princeton University, the conservation achieved by providing homeowners with daily weather-corrected feedback was 10%. The problem

involved in this initial study was that procedures for information dissemination were not practical on a large scale.

Subsequently, two research programs were developed and implemented. The first was a pilot test with a local Washington, D.C. utility using the only commercially available feedback device (Fitch Energy Monitor) on the market at the time. The Fitch monitor provides a display in cents-per-hour of total household electricity use. It also serves as a conventional digital clock, alternating the two types of information. Analysis of data from all all-electric homes outfitted with the monitor (prior to the pilot test) indicated energy savings of 12%. A representative sample of utility customers were taken and assigned at random to treatment and control conditions (70 subjects in each group). This pretest-posttest control group design will run one year concluding December 1978.

This first test was not designed to answer the question of the most effective type of feedback (DOE does not, for example, endorse the Fitch monitor), but is being used as a preliminary study advancing knowledge, especially as it relates to technical and legal dimensions, utility role, and customer acceptance. A concurrent test, involving focus groups and a controlled experiment, was designed to address issues of information effectiveness (e.g., information types, amount of information utilized, cumulative versus instantaneous, etc.)

The output from these studies will be used to develop specifications for a feedback device to be used in a major demonstration starting in the Fall of 1979. The device plus educational information will be tested in a series of experiments across geographic regions, fuel types, and fuel prices. Dependent measures will include consumption, attitudes, knowledge, and awareness of energy related factors in the environment.

#### Energy Cost of Ownership: A Communications Program

The premise for the Energy Cost of Ownership Program (ECO) is that most consumers are not yet considering energy costs in the purchase of many products. This has resulted in a major barrier to acceptance of new energy efficient technologies since many of these products have higher first costs or replacement costs.

In making a decision about which product to purchase, it is generally true that the energy saving product (e.g., insulation) or energy efficient appliance (e.g., refrigerator will power switch) will be a good investment for the consumer over the life of the product, even though it may cost more initially. Consequently, the objective of ECO is to accelerate the acceptance of energy efficient products by testing methods of increasing consumer awareness, attitudes toward, and use of the concept of energy cost of ownership in the marketplace.

The ECO Program is designed to include three major steps with consumer research in integral part in developing each step as well as evaluating the feasibility of going on to each successive step. First, an integrated marketing/communications program was developed as a pilot study in Denver, Colorado. The two major components were: (1) a communications project including paid multi-media advertising, a home energy retrofit contest, and a shopping center display of a home energy use simulator; and (2) a research phase involving measurement of the overall program and concept and evaluation of the various strategies. Advertising was done on radio and T.V. emphasizing specific products (automatic set-back thermostats, electric ignition pilot lights, storm windows). The campaign theme was "Products That Save Energy Pay For Themselves." This was displayed in all advertising and promotions and on color in-store

materials.

Utilizing a nonequivalent control group design, results were encouraging. There was a favorable impact on consumer knowledge and attitudes toward energy conservation dimensions, energy efficient products, and about costs and savings resulting from various energy measures. In addition, over 100 retailers participated in the program.

The results of the pilot test have served as input into ECO II, a major demonstration designed to create a similar positive selling environment. In this phase, even greater emphasis will be placed on the concept of energy costs and specific dollar savings. The focus, through T.V. and prints ads, will be on applying the concepts across a package of products instead of a few specific ones. This demonstration will take place with heavy private sector participation in five test markets coinciding with the market territories of the major retailers. This will allow for close evaluation of the program across varied geographic regions, fuel types and prices, and climatic conditions in order to determine the feasibility of ECO III national program the following year.

#### Summary

Each speaker offered a different perspective on the energy situation and methods for dealings with it. Millstein provided descriptive data on the current situation and changes that have occurred over the last five years. He also cited several sources for such data which can be accessed by writing him. His talk provided the background for a discussion of specific programs in Canada and the U.S. Kelly presented a broad coordinated program designed to reach a variety of populations through a phased communications strategy. Hutton focused on a specific area of conservation activity in the DOE, describing programs in which consumer research was an integral component. For further information on these programs and available materials, the appendix provides addresses and phone numbers of the speakers.

#### Appendix

Brian G. Kelly  
Energy, Mines, and Resources Canada  
580, rue Booth  
Ottawa, Canada K1A 0E4  
(613) 992-9294

Jeffrey Millstein  
Office of the Assistant Secretary  
for Conservation and Solar Applications  
Department of Energy  
1200 Pennsylvania Ave.  
Washington, D.C. 20461  
(202) 566-7263

Bruce Hutton  
College of Business Administration  
University of Denver  
Denver, Colorado 80208  
(303) 753-3316



OVERVIEW OF WORKSHOP "CONSUMER RESEARCH GOES TO WASHINGTON" <sup>1</sup>

J. J. Persensky, Chairman  
Center for Consumer Product Technology  
National Bureau of Standards

Speakers:

J. J. Persensky, Chairman

Charles Handy  
Food Economics  
U. S. Department of Agriculture

Dennis McNeill  
Bureau of Consumer Protection  
Federal Trade Commission

Richard L. Oliver  
Division of Drug Advertising  
Food and Drug Administration

Ann M. Ramey-Smith  
Center for Consumer Product Technology  
National Bureau of Standards

Raymond C. Stokes  
Director - Division of Consumer Studies  
Food and Drug Administration

Selwyn M. Waingrow  
Office on Smoking and Health  
U. S. Department of Health, Education, and Welfare

Introduction

The purpose of this workshop was to discuss the kinds of consumer research performed by Government agencies and to designate areas of research where more work is needed. Items addressed included program descriptions, indications of who actually performed the research and policy issues and problems. There are no papers describing research findings. Not all Government consumer programs were included. The objective was to familiarize researchers with some current Government programs and to stimulate the Association for Consumer Research community to think about and perhaps pursue consumer research which may have application to Government policy or research objectives.

This paper is the Chairman's summary and interpretation of the proceedings at the workshop. The paper is organized to briefly describe relevant agency programs and includes suggestions for research needs.

Current Government Programs

U. S. Department of Agriculture (U.S.D.A.)

The U.S.D.A. has a long history of consumer research supporting their role in which the consumer is viewed as the end of the food chain. Research has been centered around the creation of food policy through studies of the social science, technical, and physical aspects of human nutrition. Economic and statistical policy analysis have also been employed. Current research has been focused on four programs: the situation in the food sector; basic research on the factors affecting food choice; food policy analysis; and food assistance programs.

The food sector situation is determined through studies

of consumption trends and food purchasing behavior, with emphasis on various demographics. The National Household Food Consumption Survey is the main tool used to support this program. It provides a new, rich source of cross-sectional data and provides a data base both for U.S.D.A. and other consumer researchers.

Basic research is undertaken to determine the factors which affect food choice behavior. The factors of interest include: family size and composition; advertisements and information; health status; and, attitudes, opinions, preferences, and life styles. Other basic research addresses the economics of human nutrition, the effectiveness of nutrition education and studies of psychographic variables in shopping and purchase behavior.

Food policy analysis has, in the past, been directed primarily toward the agricultural sector. New emphasis is being placed on research in the consumer sector related to food safety and regulation as regards chemicals and additives in food and the effects of changing lifestyles with an emphasis on nutrition.

Food assistance programs are being studied through the National School Lunch Program. Again the emphasis is on nutrition with the interactive effects of the attitudes of students, faculty, and parents as a measure.

Federal Trade Commission (FTC)

The FTC has had a consumer research program since 1971 where researchers have provided consumer input to Commission staff and officials. The FTC has sponsored research in many areas including appliance energy labeling, childrens' programming/advertising, warranties, ad substantiation, and corrective advertising. FTC is also required to perform evaluative research to determine the impact of their programs and regulations. The Proceedings of the 1978 American Marketing Association Educators' Conference includes a paper by M. Mazis and D. McNeill which describes the range of programs within FTC. Two topics of specific interest are guidelines for Endorsement Advertising and Comparison Advertising. These guidelines are in the form of policy statements which are an interpretation of the Administrative Law under which FTC is currently governed.

The FTC has provided guidelines for advertisers when endorsers are used to promote products. The guidelines provide that the honest view of the endorser is expressed, that there is no distortion of the endorser's opinion or experience and that if the endorsers purport to use the products they actually do. Several research questions arise which could test the effects of the endorsement policy. Included are criteria for defining endorsers, visual effects, organizational endorsers, and effects on the consumer.

The second topic is comparison advertising. Attempts have been made to measure the broad effects of comparative advertisements but directed research is still needed to guide future policy. The questions of deceptive effects, reality of the test situation, extent of information, and effects of full disclosure of results need to be answered.

## Food and Drug Administration (FDA) - Bureau of Drugs

The Bureau of Drugs of the FDA is concerned with the regulation of the drug industry. The focus of their research has been on drug advertising and drug labeling.

The FDA is concerned with monitoring drug advertising to determine the validity of drug advertising claims and the effect of drug promotions on physician's perception of deception in advertising. Present research efforts center around physicians' behavior in responding to prescription drug promotions. The drug information included in promotional materials is evaluated for its compliance with FDA regulations. The FDA's monitoring of drug labeling focuses on the labeling of drugs as it affects both physicians' and consumers' behavior. Of special interest is the problem of including drug reactions in information which accompanies the products. Current research is aimed at developing and evaluating Patient Package Inserts for use by consumers to include information concerning drug-drug and drug-food interactions.

The FDA Bureau of Drugs is also concerned with evaluating the effectiveness of current drug advertising regulations. Legal guidelines for ads directed at physicians require that certain minimal requirements be met for providing information on which physicians make decisions. Research is performed to determine the content and format which best communicates the message.

## National Bureau of Standards (NBS)

The mission of the Center for Consumer Product Technology (CCPT) at the NBS is, essentially, to develop standardized test methods for consumer products and public safety equipment. The CCPT has no regulatory power; its function is to provide technical assistance to other government bodies with regulatory authority and to industry for voluntary standardization. There are four program areas addressed by the CCPT: Public Safety, Product Safety, Energy, and Product Performance.

The Public Safety program is represented by the work of the Law Enforcement Standards Laboratory which develops guidelines for law enforcement equipment standards. Work has been performed on items such as helmets, lights and sirens, protective clothing, and psychological deterrents to nuclear theft. Product safety work is performed primarily for the Consumer Product Safety Commission and has involved studies of bicycles, lawn mowers, hot surfaces, playground equipment, and electrical wiring.

The NBS led the introduction of the concept of appliance energy use labeling through a voluntary program initiated in 1973. The Department of Energy and the Federal Trade Commission are now leading such a program on a mandatory basis. CCPT has provided significant input to those agencies related to energy use test methods which incorporate the effect of consumers' behavior.

The Product Performance program has addressed the concept of product life-cycle performance/cost, consumer information needs, and the needs of special user groups such as the handicapped and elderly. A primary thrust was the Voluntary Consumer Product Information Labeling Program. The objective of this program was to provide point-of-sale informative labels which were designed to allow consumers to include performance information in their purchase choice.

## Food and Drug Administration (FDA) - Bureau of Foods

The Bureau of Foods of the FDA is active in consumer research to analyze consumer responses to and needs for information on foods and cosmetics. Current research efforts are devoted to consumers' understandings and use of food labeling information such as ingredients, nutrition, additives, freshness dating, and product weight.

These efforts are aimed at providing information to consumers so as to maximize its usefulness and understandability.

In addition to their role in traditional consumer research, FDA has recently initiated research-oriented surveillance of food products in order to assess the nutrition profile and other attributes of various major food categories. The goal of this research is to provide current information on food contents such as artificial flavors or colors and preservatives. Compliance oriented surveillance has been the Bureau's function since its inception. The regulation of the food industry makes use of economic cost/benefit studies which allow the FDA to study the cost impact of proposed regulations of the food industry.

The FDA also recognizes an obligation toward consumer education, especially as they make significant changes in food labeling - an effort to assess food consumption patterns is designed to measure long term patterns of eating behavior in order to address overconsumption trends of certain food categories. Assessment of consumption patterns may be used to determine what information is necessary to allow consumers with special needs to avoid certain food categories.

In addition to traditional research methods, the Bureau of Foods relies on, and is required to respond to comments from the public and the food industry at public hearings and through responses to Federal Register Notices.

## Office on Smoking and Health (OSH)

The Office on Smoking and Health is a new agency whose predecessor was the National Clearing House on Smoking and Health. The mission of OSH is to administer a national program to inform the public on the dangers of smoking, to reduce deaths and disabilities due to smoking, and to promote and stimulate behavioral, biological, epidemiological, and toxicological research on the causes and effects of smoking. OSH works with the Office of Health Information and Health Promotion to carry out their mission and to coordinate all Public Health Service information programs.

Past research has emphasized questions related to attitudes, beliefs, and practices regarding smoking among teenagers and adults, surveys of health professionals regarding their personal smoking habits and the effects on client interaction, cessation methods, and evaluation of health education both in terms of community and special populations.

A major policy question has been to define the role of the government in protecting the individual in a private behavior. Questions have included privacy, the cost to society, the differences between roles of being helpful or harassing, and differences between information/education and persuasion/manipulation.

## Implications for Future Research

The purpose of this section is to provide some insight into possible research topics of interest to members of the Association for Consumer Research. The results of such research would be helpful to Government consumer programs. This section is a synopsis of the suggestions made by the various speakers categorized by research area rather than by agency. James Bradley of the Human Factors Division of the Consumer Product Safety Commission was unable to attend the conference but did provide input to this section. The general idea behind these suggestions is to stimulate research which will guide policy rather than research which is reactive to policy decisions. Funds may be available from the agencies par-

ticipating, but it is hoped that independent research in these areas will be performed regardless of funding.

The research topics may be specific to agency missions, such as smoking cessation methods, but most are generic in nature and could be applied to nutrition and foods, advertising, drugs, safety, or product performance. Three major areas are suggested: Consumer Information, Advertising and Promotion, and Consumer Behavior. A fourth area, the Role of Government, is listed separately. Many of the research topics are expressed as questions.

#### Consumer Information

Included in this area are labels, brochures, Patient Package Inserts, and other brief information modes. Such information may be intended for consumer use at pre-purchase, point-of-purchase, or during product use.

- Valid methods, other than self-report, should be developed to determine what information consumers want, need, and will use.
- What are the criteria for determining the technical level to be used in the information?
- Should information be presented in terms of contents or performance, or should the information be provided for avoidance or to express hazards present?
- Does price or brand preference outweigh performance or safety information?
- How much information can be presented prior to sale or at point-of-sale?
- How should this information best be presented to be attracting and useable?
- What media do consumers attend to most for the purpose of gathering product or service information?
- Should specific methods be developed for gathering input on needs from special user groups, such as the elderly, handicapped, teenagers, physicians, or children?
- What presentation media are most effective for the various special user groups?
- How effective are use and care manuals?
- How can consumer information processing theory be applied to policy making?

#### Advertising and Promotion

Research topics separate from consumer information are listed for ads and promotions because of their wide usage and the special needs in these areas.

- Do ads encourage brand choice only or do they encourage generic usage?
- What are the effects of endorsers on product choice or decision criteria?
- What are the effects of organizational endorsements?
- Do visual techniques enhance the effect of endorsers?
- In what form is corrective advertisement most effective?
- How is deceptive or distorted advertisement best corrected or disclosed?
- Can consumers determine if tests in comparative advertising are appropriate to normal use? What are the effects of such presentations?
- What are the most effective modes for ads and corrective advertisements for special user groups?
- What are the effects of sales personnel and retailers on special populations?

#### Consumer Behavior

The research topics listed refer mostly to product or service use rather than pre- or point-of-purchase behavior.

- Valid methods should be developed to measure use and satisfaction with products and services.
- Normal and stereotypical practices regarding nutrition, eating behavior, smoking behavior, product use, and safety practices should be measured.
- Does the mass media effect life-style changes?
- Is mass media promotion more effective than specific consumer education techniques in altering behavior?
- What are the economic factors related to nutrition?
- How does risk-taking interact with safety practices?
- What methods can best bring about cessation of dangerous practices, e.g. smoking or unsafe use of products?
- What are the interactive and feedback effects on consumer behavior of the attitudes of families, physicians, patients, and other professionals.

#### Role of Government

Many problems are being addressed by the Government in the consumer area. Are they all appropriate to Government intervention?

- What limits should be placed on consumer related regulation?
- Do nutrition, safety, smoking, or advertising regulations infringe on the privacy, freedom of choice of the individual?
- What is the societal cost of consumer problems and regulation?
- Is safety an important concern of consumers?
- Should the Government attempt to modify consumer behavior?

A final question is posed to the members of the Association of Consumer Research. Should you become an active force in strengthening Government consumer programs by performing policy related research regardless of funding or by responding to Federal Register requests for comment? Remember that not all Government consumer programs were represented in this presentation. Any of the speakers could prove to be valuable resources when planning a research project.

\* \* \*

<sup>1</sup>Contribution of the National Bureau of Standards. Not subject to copyright.

W. Fred van Raaij, Tilburg University, The Netherlands

It seems impossible to generalize over the kinds of consumer research that is carried out in various European countries. Many structural differences exist between the economic systems in Western and Eastern Europe. Other differences are observed between the Northern and the Southern European countries. While in Eastern Europe all market research is carried out by public agencies, in Western Europe, both public and private agencies carry out market and consumer research.

Flemming Hansen (Copenhagen School of Business Administration and Economics, Denmark) described the commercial consumer research carried out in Western Europe, mainly by private agencies. The number of market researchers in the various countries can be estimated from the ESOMAR membership directory. The United Kingdom (75) and West-Germany (68) have most ESOMAR members, followed by The Netherlands (42), France (40), Italy (35), and Switzerland (33). Spain (17), Denmark (16), Belgium (15), Finland (12), and Norway (11) have a lower number of members.

The total world market for market research in 1977 is 735 M£ (million pounds). In North-America, 340 M£ is expended; in Europe, 275 M£. Of this amount, the turnover in West-Germany is 87 M£; in the United Kingdom 52 M£; in France 51 M£; and in The Netherlands 20.4 M£.

The users of market and consumer research in Europe are manufacturing companies (28%), food, drink, and tobacco companies (22%), advertising agencies (15%), academic institutions (10%), and services/trade (6%). Noncommercial research is of increasing importance. Sponsors of noncommercial research are the European Community (EEC) in Brussels, national and local governments, public and academic institutes, media, and private foundations (e.g. Volkswagen Foundation).

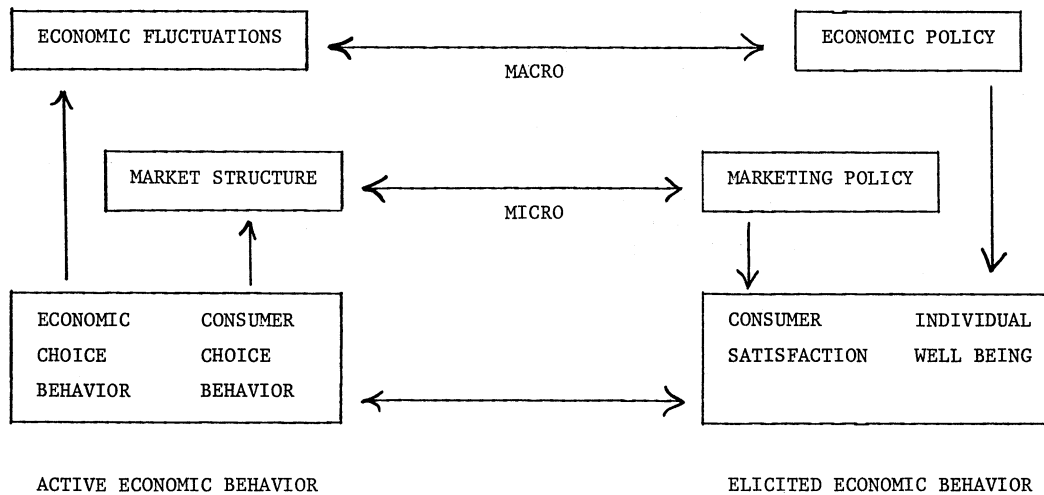
Comparing the costs of market research in the European countries, the United Kingdom offers the lowest prices for market research. Tremendous differences are found between countries and between agencies within countries. France, West-Germany, Sweden, and Switzerland are expensive countries to carry out market research. Intermediate price levels are found in Italy, Belgium, Denmark, and The Netherlands.

Susan Douglas (CESA, France; now New York University) described the consumer research carried out in academic institutions. Compared with the U.S.A., not very much consumer research is carried out at European universities. The emphasis is on consumerism research and research for consumer policy. Factors that seem to block the full development of consumer research are the lack of funding and the communication (language) problems between countries. The emphasis in training is more on philosophy of science and less on techniques and methodology. In the various countries, some academic institutes conduct research on an international level and publish reports in English. Some researchers, especially in France, received their training in the U.S.A. Many publications written in a national language remain unknown to researchers in other countries.

The European Institute for Advanced Studies in Management in Brussels plays an important role in creating a network of researchers and a communications center. This institute established the European Academy for Advanced Research in Marketing, which has nearly 400 members from twenty countries, including the U.S.A. and Canada. The next annual conference of the Academy will be held in Bronigen, The Netherlands, April 10-12, 1979.

The Annual ESOMAR conference will be held in Brussels, Belgium, September 2-6, 1979 with the theme "The Chal-

Scheme 1: The micro and macro circuits



This scheme was developed by the group of economic psychologists of Tilburg University: Gery M. van Veldhoven, Theo M.M. Verhallen, and W. Fred van Raaij.

lence of the Eighties."

While the Academy is a typical marketing research organization, the European chapter of the Association for Consumer Research will coordinate and stimulate consumer research from the academic, industrial, and consumer protection perspective.

Fred van Raaij (University of Tilburg, The Netherlands) described the trend in Europe from micro consumer research to macro consumer research, or from research for specific brands to research for government and public institutions. Scheme 1 gives two circuits, the micro (inner) circuit and the macro (outer) circuit. These circuits are similar in structure. Consumer choice behavior (specific brand choice, brand loyalty, innovative behavior) influences the market structure of brand shares, market composition, competition, and substitution. Economic choice behavior forms the macro circuit (generic product choice, spending vs. saving decisions, investment, demand) influences the economic situation of a country, inflation, and economic growth.

Marketers react to their perception of the market structure by implementing marketing mix factors in order to increase their market share, consumer satisfaction, and their company's profit. Economic policy makers react to the economic fluctuations, using taxation, interest rates, wage policies in order to maintain or increase citizen's well-being.

The emphasis in Europe tends to shift to the macro circuit leading to research on consumer well-being and satisfaction, consumer complaining behavior, the environmental effects of consumption, and research on poverty. This research is carried out on a cross-national level by the European Community (EEC): "The European Consumer" (1976) and "The Perception of Poverty" (1977), and by the Organization for Economic Cooperation and Development (OECD), mainly on social indicators. On a national level, the proportion of government-sponsored research increases in most European countries. Governmental agencies tend to give priority to political problems over consumer problems to be researched on a large scale. In many countries, social indicators measuring well-being or quality of life, even including psychosomatic indicators, are developed.

#### References

"The European Consumer: His Interests, Aspirations, and Information," Brussels: Commission of the European Community, Wetstraat, 1049 Brussels, Belgium. May 1976.

"The Perception of Poverty in Europe," Brussels: Commission of the European Community. March 1977.

#### Useful Addresses

European Institute for Advanced Studies in Management, Secretariat: Ms. Gerry Dirickx, Stephanieplein 20, 1050 Brussels, Belgium.

European Society for Opinion and Marketing Research (ESOMAR), Secretariat: Raadhuisstraat 15, 1016 DB Amsterdam, The Netherlands.

#### European Journals

European Journal of Marketing, MCB Publications, 198/200 Keighly Road, Bradford, West Yorkshire, England BD9 4JQ.

European Research (ESOMAR), Editor European Research, Kluwer, P. O. Box 23, Deventer, The Netherlands

Journal of Consumer Policy, Associate Editor Heiner Imkamp, University Hohenheim, P. O. Box 106, 7000 Stuttgart 70, West-Germany.

Journal of the Market Research Society, The Market Research Society, 15 Belgrave Square, London SW1X 8PF, U.K.

Market Research Abstracts, The Market Research Society, 15 Belgrave Square, London SW1X 8PF, U.K.

## PRAGMATIC APPLICATIONS OF CONSUMER RESEARCH IN RETAILING

Charles W. King, Purdue University  
Douglas J. Tigert, University of Toronto  
Lawrence J. Ring, University of Virginia

### Abstract

Pragmatic consumer research in retailing has burgeoned in recent years based on leadership by sophisticated consumer researchers in savvy retailing organizations. This session discusses pragmatic consumer research in retailing from the vantage points of merchandising management and retail researchers across divergent retailing contexts. Summary generalizations and new directions for future research are reviewed.

### Introduction

In recent years, the retailing community has developed increased sophistication in using consumer research. The objective of the session, Pragmatic Applications of Consumer Research in Retailing, was to increase dialogue between the retailing community and the academic community interested in consumer research in retailing.

More specifically, the session was organized to present some practical applications of consumer research in retailing with managerial implications and to identify major developments and new directions for consumer research in the broad retailing arena. Toward that goal, a diverse set of speakers was recruited. The speakers included merchandising management, retailing researchers, and academics involved in consulting and more theoretical consumer research in retailing. The speakers from the retailing community were:

Mr. Vern L. Page--National Merchandise Manager of the Bath, Slumber and Hostess Products Department, Sears Roebuck and Company

Mr. Ralph D. Holt--Research Director, Federated Department Stores

Mr. Alfred L. Morin--Associate Research Director, Burger King Corporation

The academic research team of King, Tigert and Ring developed a backdrop of managerial research in retailing to place the pragmatic applications in perspective, synthesized the session presentations and discussion into this summary statement, and highlighted major conclusions and future consumer research directions in retailing.

### Managerial Research in Retailing: A Perspective

In recent years, there has been a marketing research explosion in the retailing community, particularly among the major retailing organizations. Hundreds of articles, research reports and research notes/observations/assertions appear annually in the retailing trade press, in the industry trade channels, and in the academic literature.

The growth in retailing research, however, is a comparatively recent development. Historically, retailing has been one of the last outposts of the "gut feel/learn-at-the-retail counter" management school. And, within some departments of the major retailers today, and among most of the smaller retailing firms, the "research revolution" notwithstanding, that management style still prevails. In fact, much of the "theory" of contemporary retailing

is a compendium of "philosophical commentary," "sacred cows," and folklore based axioms/principles that are accepted as "conventional wisdom" among operationally savvy retailing managers.

To complicate the problem further, there is no actual "retailing industry." Rather, there is a series of different types of retailing institutions or sectors of retailing, e.g., the supermarket sector, the general merchandise store field, the general merchandise discount store world, the specialty/boutique retailing milieu, the service retailing area, the special product merchant, e.g., the fast food retailer, the lawn and garden retailer, the gasoline retailer, etc. Each of these sectors has its own operating philosophy and perceptions of consumer behavior and competitive reality.

The increase in managerial interest in research has been the product of rather dramatic changes and challenges in the competitive environment of contemporary retailing. The retailing research thrust has come from several sources. Retailing research is generated within the retailing organizations to analyse daily operations, to track sales and profit performance, to monitor store imagery among consumers and to track the changing competitive market structure. Likewise, the volume and potential sophistication of internal sales analysis now available to the retailer is almost unlimited given modern computer technology and innovations in retailing operations such as the point-of-sale accounting system.

Major manufacturing resources conduct research to support their products' sales through the retailing distribution channel. The mass media are increasingly researching consumer shopping and buying behavior to support and solicit the retailers' advertising dollars.<sup>1</sup> Likewise in academia, Speh (1978) notes that the retail environment has become an important research setting for a number of academic disciplines.

Admittedly, much of this retailing "research" may be methodologically unsophisticated by the criteria of the marketing analytic and the "scientific method." The content, however, is part of the burgeoning information "fog" in which the retailing manager operates.

King and King (1978) surveyed and synthesized the trade and academic literature, interviewed researchers in major retailing organizations, and identified a series of research themes or major topic areas that characterize the emerging retailing research literature. As presented in Figure 1, the focus of retailing commentary/research is very broad and multi-dimensional.<sup>2</sup>

<sup>1</sup>The CHICAGO SUN-TIMES, in conjunction with King and Ring, is currently developing a massive consumer research data bank on male and female fashion involvement, attitudes, buying behavior, communications exposure, and consumers' perceptions of leading fashion retailers' images in the competitive fashion market in Chicago to support the Chicago retailing community and its advertising planning.

<sup>2</sup>For a discussion of on-going research directions within the retailing community, see Pessemier (1978).

FIGURE 1

MANAGERIAL RESEARCH IN RETAILING: SOME MAJOR TOPICS

1. The Dynamics of Consumer Retail Attitudes and Shopping/Buying Behavior
2. Consumer Information Processing Research
3. Store Image Research
4. Pricing Research
5. Store Location Research
6. Mathematical Modeling
7. Retailing Personnel Selection, Training and Management Development
8. Developments in Retail Distribution Channels
9. Retail Operations Management and Performance Analysis and Control
10. Public Policy, Government Relations and Social Responsibility
11. General Marketing and Retailing References
12. Periodic, Routine Market Position Measurement and Repeat Monitoring

Retail Merchandising Management and Consumer Research

Vern L. Page, Sears, Roebuck and Company, dealt with consumer research as a pragmatic input to the marketing planning and merchandising function of the merchandising manager. Toward that goal, the discussion outlined the Sears organization and the buying/merchandising function, the role of consumer research in that system, the major dimensions of consumer research analysis, the activities of competitive monitoring and the measurement of retail market effectiveness.

The Sears Organizational Structure and the Buying Function

To place the role of consumer research in perspective, the Sears organization structure and the internal buying function were reviewed.

Sears is organized around two operating units: the field organization and the headquarters organization. The field organization is comprised of fourteen catalog merchandise distribution centers, 2900 catalogue sales facilities and 862 retail stores. The field organization is the selling arm of Sears.

The headquarters organization consists of the corporate executive group, a series of support functions including consumer research and the corporate buying activity. The headquarters organization plans, selects, and buys the merchandise which the field organization sells.

Within headquarters, the buying activity is performed by some fifty buying departments, each with responsibility for a specific set of products. Each buying department is managed by a National Merchandise Manager and features a catalog and retail marketing staff as well as a series of buyers.

The Bath, Slumber and Hostess Products Department, the context upon which this presentation is based, is composed of ten buyers covering the entire range of products in this category, including towels, bath

accessories, bath carpet, shower curtains, comforter pillows, blankets, sheets, mattress pads, hostess products, etc. The Sears buyer is responsible for reviewing, developing, selecting, pricing, buying and servicing all of the products sold in the category seasonally or annually depending on the products in question. In this department, the product lines are changed annually.

The Marketing Plan and Consumer Research

The buying function is managed through preparation of a written marketing plan from which is developed the merchandising plan for the department. The first step in preparing the marketing plan is the development of a comprehensive situation analysis of information about the consumer, the competition, the market structure, the product and the manufacturing resource situation. From this data, the problems and opportunities are identified within the product line. Based on that analysis, a series of line performance objectives are defined and the tactical plans are prepared to maximize the line's performance.

Consumer research input into these marketing and merchandising plans comes from two broad research streams: on-going, routine tracking research which generates competitive market structure detail across and within product categories and special, focused research projects addressing a particular product or operating problem. More specifically, at Sears, the research data are produced from these sources:

- 1) "Databank"--a Sears developed research program that examines the home fashions product purchasing behavior of 20,000 United States households by product line annually.
- 2) The "Major Market Studies"--a series of products important to Sears are evaluated and tracked in each of the major retailing markets in terms of market share and competitive positioning.
- 3) Qualitative and quantitative "special studies" periodically conducted to focus on specific aspects of the product line and market conditions, as needed by the buyer.

Major Dimension of Consumer Research Analysis

Consumer profile analysis, particularly in terms of demographic characteristics, is an important area of retailing research. More specifically, consumers are measured in terms of where they live, annual family income, age of the female head of household, and occupation of the male head of the household. Analysis centers on profiling Sears customers versus competition and on tracking of geographic sources of sales volume and that of competition.

Because the characteristics of the housing unit are a critical element in the sale of bath, slumber and hostess products, data are collected about home features, e.g., whether the home is rented or owned, the style of the house, and the number of bathrooms, bedrooms, dining areas, etc. These data are then used for better understanding consumer buying behavior in the product category.

To better direct marketing communications, the media habits in terms of television viewership, radio listenership, and magazine usage are monitored for Sears customers versus competition's customers.

Additionally, the buyer monitors a variety of other dimensions of consumer buying behavior such as:

- 1) The annual purchase incidence of the product to measure and forecast market potential.
- 2) Total number of items in each product area that are purchased annually.
- 3) Gift or self consumption buying.
- 4) Purchase at sale or regular price.
- 5) The month in which purchase is made to detect seasonality of buying behavior and merchandising opportunities.
- 6) Coordinated products that are purchased at the same time or are integrated in the purchase dynamic.
- 7) The average price paid.
- 8) The identity and profile characteristics of the buyer/user where available.

#### Competitive Monitoring and Measuring "Retail Market Effectiveness"

Market structure analysis and tracking of movements by competitors in the market are important activities in consumer research. Share of market, Sears versus its major competitors, is an extremely important point of analysis. Share of market is calculated by total product category and by sub-categories within the broad product group. Likewise, tracking the shopper's shopping patterns in competitors' stores gives insight into the products and departments where the customer is doing comparative shopping and the competitive products that must be faced.

An important consumer research tool centers on measuring "retail market effectiveness." The measure of "retail market effectiveness" revolves around the issues of "drawing power," getting customers to shop the particular retail outlet for the product category, and "selling power," the conversion of the shopper into an actual retail buyer of the product. Comparative monitoring of the "drawing" and "selling" power of Sears versus competition can measure retail selling effectiveness.

#### Integrated Research Programming and Marketing Planning

In addition to the routine consumer monitoring and tracking research, Page presented several case studies in the areas of blankets and pillows. The case studies described integrated research programs which took the product from initial concept definition, through product development, into commercial introduction, followed by measurement of market penetration and retail market effectiveness.

#### Consumer Research in Retailing: A Perspective on Where We Are and Where We Are Going

Ralph D. Holt, Federated Department Stores, approached consumer research from the vantage point of the sophisticated consumer researcher who must integrate the research contribution into the retailer's decision making process. Toward that goal, the discussion outlined the consumer research function within the Federated organization, presented the major consumer research activities, discussed conceptual approaches to segmentation and store image analysis, and outlined some future directions for research development.

#### The Federated Department Stores Organization

The Federated Department Stores organization is composed of fifteen relatively autonomous department store

divisions representing some of the bellwether department stores in the United States. The group includes such leaders as Abraham & Straus, Bloomingdale's, Bullock's, Filene's, Lazarus and Rich's. Additionally, Federated is diversified into women's speciality apparel with I. Magnin and Bullock's Wilshire, hardgoods with Gold Triangle, discount chains with Gold Circle and Richway and supermarkets with Ralph's. Though diversified, the major share of the firm's approximately five billion dollars in retail sales is accounted for by the department store divisions.

A characteristic of the Federated organization, the divisions operate as very autonomous staff-lean management units. Each division operates, on the average, six to seven retail stores. The corporate offices located in Cincinnati, Ohio, houses the basic corporate functions. The staff departments such as the Consumer Research Department operate as consultants to the divisions and to top corporate management.

#### The Role and Scope of the Consumer Research Department

The Consumer Research Department at Federated Department Stores is 15 years old and specializes in survey research and marketing planning. The staff totals thirteen people, eight of whom are professionals. The group conducts approximately 100 projects each year.

The role of the Consumer Research Department is to serve as an operating problem answer generator. The management style of the Federated operating executives is very "action focused" with a pragmatic, "bottom line" orientation.

The challenge to the marketing research function is straightforward. Research must be problem focused. Data must be clear and conclusions specific. The implications must be "actionable" in terms of the specific tactical actions that the operating management can execute now.

The scope of the research effort includes the following types of research activities:

- 1) Development of sales estimates for potential stores to support real estate decisions and financial/marketing planning in conjunction with another corporate staff group, the Area Research Department.
- 2) Identification of growth opportunities for the existing stores as input to marketing planning.
- 3) Operating information generation in answer to specific operating questions, e.g., customer reaction to sales service.
- 4) Measurement of employee attitudes toward their jobs, management, new compensation and benefit programs, etc.
- 5) Miscellaneous problem solving as required.

#### The Development of Sales Estimates for Potential New Stores

A major activity of the Consumer Research Department is the generation of information for use by the Area Research Department in building sales estimates for potential new stores to guide expansion decisions. This research area reflects the breadth of consumer research at Federated.

Developing sales estimates for potential new stores build upon measurement of two basic concepts of market assessment: trading area determination and market quality.



In defining the trading area of a potential store site, the first step is the measurement of geographical trading areas through "point of origin" studies. These "P-O-O" studies are intended to identify the trading area of a given shopping center by simply intercepting people in the mall to define their home addresses.

Given alternative retail store sites and the travel patterns of the respective target consumers, the relative overlap of alternative sites can be measured in terms of their geographical proximity to the consumers' shopping/travel patterns. The pragmatic store location question is, which alternative site would be the best location to reach the target market?

Market quality assessment is aimed at measuring the "trading quality" of the geographic market. Market quality assessment involves a number of different inter-related analyses focusing on:

- 1) Demographic data update. Usually, particularly in rapidly changing growth markets, available secondary Census data are old--frequently several years out of date--and do not reflect current market expansions and population movements.
- 2) The "quality level" of consumer shopping patterns. Do consumers in the market tend to buy budget priced goods, medium priced goods, or better/higher priced goods?
- 3) Fashion involvement. What is the level of fashion involvement of consumers in the market area. To what extent are consumers interested in buying the very latest fashions?
- 4) Current shopping behavior of consumers in the geographical market on a store-by-store comparison for specific merchandise categories. Analysis frequently focuses on the "store shopped most often/the headquarters store" compared with other stores across product groups.

#### Some Conceptual Approaches in Retail Market Segmentation

The on-going consumer research at Federated has identified two basic dimensions of retail market segmentation that appear particularly important to the general merchandise product category. The two basic dimensions are "quality acceptance" and fashion involvement.

#### "Quality Acceptance"

The "quality acceptance" measure refers to the consumer's general propensity to buy lower priced goods, medium priced goods, or higher priced goods. The concept builds upon the following propositions that have emerged from empirical research across a variety of studies:

- 1) Store shopping behavior is related to the price paid for individual items. Higher priced goods tend to be bought from department and speciality stores; lower priced goods tend to be bought from discount stores.
- 2) Shoppers tend to be consistent in their propensity to buy lower priced or higher priced merchandise across product categories. People who tend to do purchase on the higher end of the price spectrum for one kind of merchandise tend to be on the higher priced end for other kinds of merchandise.
- 3) The "quality acceptance" measure is a better indicator/discriminator of shopping behavior at particular types of stores and at individual retailers than traditional demographic measures.

The "Quality acceptance" scale is used to categorize consumers into five segments based on their price point shopping styles. The local Federated store's customer share in each of the segments compared with competition is tracked as a measure of merchandising performance and target market penetration.

#### Fashion Involvement

Fashion involvement deals with the female consumer's interest in and adoption of new fashions. After extensive exploratory research, a "contingency approach" was developed to measure fashion involvement. The underlying concept was that, for the consumer to adopt new fashions, there are certain requisite conditions that must be met:

- 1) Dress size must be small, i.e., size 10 or smaller, to access the consumer to the more fashionable merchandise which is, in fact, more readily available in smaller sizes.
- 2) Consumers must spend a reasonable amount on fashion apparel annually, arbitrarily defined as \$500 per year.
- 3) Consumers must have a manifest interest in fashion based on self declared fashion interest or other surrogate variables, e.g., fashion magazine readership.

Based on the contingency approach and these variables, a scale is developed to measure "the consumer's propensity to adopt new fashions."

Using the fashion involvement scale, four segments are identified ranging from low to high fashion involvement. As with the "quality acceptance" measure, the local Federated store's customer share in each of the segments compared with competition is monitored and tracked over time as a measure of merchandising performance and target market penetration.

#### Store Preference and Image Measurement

An important element in consumer research at Federated deals with shoppers' store preferences and store images. As basic measures of shoppers' store preferences, two basic questions are used:

- 1) "Which stores do you usually shop for your own apparel and household goods..?"--used to identify "occasionally shopped" stores;
- 2) "Which one do you shop most often...?"--used to identify the consumer's "headquarters store." This measure correlates very highly with store market share ("r" in the range of .85).

In terms of measuring store images, seven basic image item questions as listed in Figure 2 are typically used. Based on factor analysis, these items are reduced to four basic image dimensions of convenience, value for the money, fashion orientation, and salespeople availability.

#### A Typical Federated Study

Building upon the foregoing discussion, a typical Federated consumer research study would involve the following analytical modules:

- 1) A current demographic profile of the particular geographic market is developed in terms of traditional demographic measures, age, education, income, occupation of chief wage earner, etc.

FIGURE 2

STORE IMAGE DIMENSIONS

<u>Basic Image Item Questions</u>	<u>Factor Concept</u>
Convenience	Convenience
Selection in your price range Value for the money	Value
Selection in styles you like Selection of latest fashions Having exciting merchandise	Fashion
Availability of sales people	Sales service

- 2) Using the consumers who "shop federated most often" as Federated's most loyal shopper group, analysis is focused on the store's penetration in each of the demographic segments.
- 3) Applying the "quality acceptance" measure to the local market, its profile on this dimension is compared with the national standard. Using the five quality acceptance segments, Federated's customer share is compared with other competitors.
- 4) Applying the fashion involvement measure and the resulting four fashion segments, Federated's customer share is compared with other competitors.
- 5) Store images of Federated compared with other competitors are measured on the seven items and four factors presented in Figure 2 for both customers who shop the local Federated store most often and for those who occassionally shop.

From this type of consumer research project comes valuable input to retailing management for defining the store's current position and its market penetration. Likewise, based on this bench mark, the impact of new merchandising strategies can be measured in terms of penetration changes in specified target market segments.

Areas for Future Development

Two basic areas for future development of consumer research in retailing have been identified. In the field of store image research, a variety of measures of store image have been used in the literature and in the Federated research efforts. Two questions arise:

- 1) Are these the correct image dimensions to be analysed and are the measurement procedures valid?
- 2) Assuming a reasonably comprehensive set of store image dimensions can be operationally defined and measured, what is the relative importance of the various image dimensions in influencing store choice and shopping behavior?

Little research has treated this very critical issue of image dimension importance which is essential to the retail manager in allocating merchandising resources.

The second critical area for development centers on the actionability of the research data. Consumer research in retailing is increasingly effective in identifying problems. The research, however, has too often been unable to produce clear solutions to the identified problems to guide the retailing manager.

An Integrated Consumer Research Program of Market Monitoring

Alfred L. Morin, Burger King Corporation, approached the consumer research task at the Burger King Corporation as an integrated on-going consumer research market monitoring function. In that context, Morin outlined the broader marketing research activity within the Burger King Corporation and presented an illustrative series of consumer research projects that comprised a comprehensive and integrated consumer research program.

The Role and Scope of Marketing Research at the Burger King Corporation

The Marketing Research Department at the Burger King Corporation is composed of the Research Director, the Associate Research Director, five Research Managers and eight analysts. The role of the group is to support the retail marketing operations with pragmatic, actionable research and operational recommendations. Toward this goal, the Marketing Research Department does research for a wide range of corporate client departments such as the Procurement Department, the Real Estate Department, the Operations Research and Development Department as well as the field marketing organization.

The emphasis on actionable research results is reflected in the funding of the Marketing Research Department. The corporate budget supports the managerial personnel, the Research Directors and the Research Managers. The corporate research client end users, however, pay all out-of-pocket research project expenses plus assigned analyst time and overhead costs associated with specific projects. Therefore, the department research must be cost-effective to secure corporate clients and to survive.

The types of research and the relative budget commitments to the various research activities are summarized in Figure 3. Consumer research commands the largest budget and is conducted both within house and through J. Walter Thompson, Burger King's advertising agency. Product testing is a unique retailing research function but is mandatory in food retailing. The operations research and the local store area research are relatively new additions to the function. They focus on the mechanical aspects of day-to-day store operation within the specific local market competitive context.

FIGURE 3

TYPES OF RESEARCH CONDUCTED AT BURGER KING CORPORATION

<u>Research Type</u>	<u>Budget</u>
Consumer Research	\$ 500,000
Sales Analysis	100,000
Product Testing	100,000
Operations Research	65,000
Local Store Research	50,000

Consumer Research at Burger King: An Integrated Program

The consumer research effort within the Marketing Research Department at Burger King has been designed as an integrated research program involving three tiers, national marketing research, geographic market wide coverage and local restaurant focused research. The research program is intended to generate a comprehensive information system ranging from macro national marketing information applicable to broad strategic planning down to local restaurant image assessment and operations evaluation for tactical outlet management. Illustrative

projects representative of the three tiered consumer research program are discussed in the following sections.

#### National Consumer Research

The national consumer research thrust involves several major research areas including:

- 1) Adult advertising awareness tracking.
- 2) Children's advertising awareness tracking.
- 3) Large scale menu census: "Who eats what" with comprehensive customer profiling based on a sample of 22 bellwether restaurants.
- 4) Strategic Segmentation Study designed to support full scale brand and product review in 1979.
- 5) Advertising copy testing using sophisticated technology developed by J. Walter Thompson.

The adult and children's advertising studies periodically track brand awareness and usage, advertising awareness and advertising recall. The Strategic Segmentation Study is an omnibus research effort including attribute and similarities analysis, importance of particular attributes in the shopping process, attitudes toward the "eating process" and the role of fast foods hamburger restaurants in that process, life style dimensions, etc.

#### Market Wide Consumer Research: "Marketrac"

The market wide consumer research monitoring project, "Marketrac" focuses on a geographic trading area, the area of dominant influence (ADI). The purpose of this type of penetration survey is to provide quantified consumer perceptions of Burger King's position within a specific ADI relative to competition, relative to Burger King's share of fast food hamburger distribution in the area, and relative to marketing, demographic, and operational strengths and weaknesses in the market.

The research is a joint venture between the regional Burger King field marketing management and the corporate Marketing Research Department. The Regional Marketing Manager develops information on Burger King's share of the total chain fast food hamburger restaurants in the market. The Marketing Research Department conducts a telephone survey of approximately 400 consumers randomly selected from the market area. The telephone survey focuses on the key dimensions of awareness, trial, "use most," "used recently/within the past 4 weeks", and ratings on specific attributes of the Burger King image and product.

The analysis focuses on:

- 1) Burger King awareness, trial, use most, and used past four weeks across various relevant market segments, e.g., adults vs. teenagers, males vs. females, heavy vs. light fast food hamburger restaurant users.
- 2) Comparison of Burger King's "use most" score to the share of units of distribution to evaluate whether Burger King is getting its "fair market share" based on its physical store density in the market.
- 3) Evaluation of marketing/operational opportunities based on the strong and weak image features reported by the "use most" the "used recently" and the non-user market segments.

#### Local Restaurant Focused Consumer Research

The local restaurant focused consumer research is aimed at providing a better understanding of consumer buying behavior in the fast food hamburger restaurant context and the competitive market structure at the local restaurant level. The output should produce actionable recommendations for the restaurant manager.

Three types of research studies are illustrative of this area:

- 1) The Trading Area Survey Kit (TASK).
- 2) "CANNIBEST".
- 3) Local Store Attitude Study (Developmental).

#### The Trading Area Survey Kit (TASK)

The trading area survey is designed to provide data on where customers are coming from and going to related to their Burger King patronage. The surveys are administered by a trained interviewer at the door of the subject restaurant to every "Nth" customer entering the restaurant.

Customers are quizzed regarding where they were prior to coming to Burger King, where they are going after leaving and where they live. Additionally, the respondents are surveyed regarding the shopping group size, composition, frequency of visiting Burger King and competitors, and opinions about the local Burger King operation.

These diagnostic trading area surveys are conducted at the discretion of the Burger King field management, but are typically conducted twice a year. In problem situations, after the problem identification/diagnosis survey is completed and remedial action is executed, post-action surveys may be taken to measure impact. These surveys constitute the basis for the six-month local restaurant marketing plan.

#### "CANNIBEST" (Cannibalization Estimation)

The purpose of "Cannibest" is to estimate the potential impact of a proposed restaurant on the retail sales of an existing Burger King restaurant.

The "Cannibest" research effort is again a joint venture. The Marketing Research Department provides information about the customers of the existing Burger King sites. From that data, the Burger King real estate representatives make an estimate of cannibalization that will result from a new operation at the proposed site. When a "Cannibest" study is requested, the Marketing Research Department sends professional interviewers to the existing restaurant. Approximately 500 customers are interviewed, typically on a Thursday and a Saturday. The interview protocol is very similar to that used in the Trading Area Surveys.

The real estate representatives take the research data and integrate trading area knowledge, traffic flow, neighborhood composition, natural and man-made trading barriers, etc. From this information bank, the Cannibalization Estimate is made to guide the expansion decision.

#### Local Store Attitude Study

The local store attitude study is a qualitative and developmental effort to generate qualitative insight into consumers' problems with specific restaurants at the local level. The research approach is an effective way for owner/operators to solicit communications from their store customers efficiently.

Methodologically, a convenience sample of any available customers or non-customers is interviewed. The interviewer is typically a restaurant crew member. The interview questionnaire is short and merely probes for specific problem areas and suggestions for improvement. The respondent's responses are recorded verbatim on the questionnaire in as much detail as practical.

The analysis involves the restaurant operator reading the comments and considering action responses. Special emphasis is placed on comments from customers who have visited the restaurant once or twice and have not returned.

#### Pragmatic Consumer Research in Retailing Some Conclusions

Building upon the foregoing presentations and experience in retailing research in both academic and consulting roles, the academic team has drawn several basic conclusions regarding consumer research in retailing and suggested directions for future research in the area.

- 1) Taking a macro-view, the retailing literature and pragmatic retailing practice indicate that retailing research can be organized around three broad themes:
  - a) Consumer shopping dynamics, decision making, choice processes, and ultimate buying behavior.
  - b) Operations analysis focusing on the day-to-day mechanics of retail institution operations, e.g., store layout, organizational structure and personnel management, merchandising, etc.
  - c) Performance measurement in financial terms, e.g., sales per square foot, return on sales or investment, and marketing terms, market share or penetration of target markets compared with competition.

Consumer research, while central to understanding consumer shopping/buying behavior, can make major contributions in the other areas of operations analysis and performance measurement as well.

- 2) Consumer research within the retailing firm should be organized into an integrated program involving a series of independent but interrelated projects which can contribute to all three major information zones.
- 3) Consumer research within the retailing institution, to secure adequate funding and to have maximum impact on retailing management, must be clear and interpretable into actionable recommendations with immediate payoff.
- 4) Consumer research proposals must be cost effective in methodological design and justified in a cost-benefit context to secure funding.
- 5) The dilemma of breadth of research coverage will continue to plague the retailing research community. Should the focus be on the stock keeping unit (SKU), e.g., "the green socks syndrome," or should research deal with the more macro issues of "the store" or "the department" versus the merchandise? Should market monitoring be aimed at broad national developments or should the focus be on local market monitoring and specific store strategies and tactics? Obviously, both macro and micro types of information are needed. But in the reality of scarce research resources, which should have priority?
- 6) Generalizations produced by consumer research in retailing will continue to be elusive because of the

pragmatism of the field and the heterogeneity of the retailing milieu. Major theoretical developments, however, can be realized when consumer research is focused on specific product contexts in specific retailing institutions, e.g., consumer fashion buying behavior within department and specialty stores or image dimension importance for fast food hamburger restaurants.

- 7) Several major consumer research thrusts in retailing have emerged in recent years and will become increasingly visible over the near term in the maturation of this pragmatic research tradition:
  - a) There will be an expansion of routine, periodic local market monitoring of consumer shop shopping/buying behavior and competitive store positioning across the various mass merchandising retail institutions of supermarkets, fast food outlets, general merchandise department and fashion speciality retailers, etc.
  - b) Consumer life style research, already well established in other marketing arenas, will become increasingly important in retail market segmentation and market structure analysis and in targeting retail marketing and merchandising programs at specific market segments.
  - c) Store image research, long intuitively appealing to and reasonably accepted by knowledgeable retailers, will move to a higher level of methodological and analytical sophistication and assume a major role in the dynamics of store positioning and market monitoring.
  - d) Consumer information seeking and processing research relating the roles of the mass media and in-store shopping in retail marketing communication will have high potential pragmatic payoff to the merchandising function.
  - e) New electronic technology, e.g., point-of-sale accounting and optical scanning equipment, will open whole new opportunities for diagnostic consumer research in retailing.

In summary, consumer research in retailing is becoming an integral part of the management information system for the informed retail manager. The outlook indicates that consumer research can make an increasing contribution to understanding consumer behavior in the retailing milieu and in supporting retailing management decision making.

#### References

- Charles W. King and Sandra E. King, "Managerial Research in Retailing: A Perspective," an unpublished working paper, 1978.
- Edgar A. Pessemier, Some Current Directions for Retailing Research, September, 1978, Institute Paper No. 676. Institute for Research in the Behavioral, Economic and Management Sciences, Krannert Graduate School of Management, Purdue University.
- Thomas W. Speh, "A Summary of Doctoral Research in Retailing," The Journal of Retailing, Summer, 1978, pp. 71-80.

OVERVIEW OF THE "HOME ECONOMICS RESEARCH:  
WHY GIVE A DAMN?" WORKSHOP

E. Thomas Garman, Virginia Polytechnic Institute and State University

Introduction

This workshop featured five prominent speakers from the field of home economics:

- E. Thomas Garman, Professor, Department of Management, Housing, and Family Development, V.P.I.
- E. Scott Maynes, Chairman, Department of Consumer Economics, Cornell University
- Earl W. Morris, Professor, Department of Family Environment, Iowa State University
- Patricia Sailor, Director, School of Home Economics, Louisiana State University
- Lucille Wakefield, Head, Department of Food and Nutrition, Florida State University.

The purpose of the workshop is spelled out in some detail in the opening remarks of Professor Garman. Following these are summary reports on some of the prospectives raised at the workshop. Citations of papers, relevant journals, and personal sources are included to assist interested members of ACR in pursuing specific topics.

Home Economics Research: Why Give a Damn?

by E. Thomas Garman

Home economists everywhere are most pleased to have the opportunity today for us to speak to members of the Association for Consumer Research regarding the importance and utilization of research conducted by home economists. The provocative title ending with "why give a damn?" seems appropriate because all too often our individual tendencies to specialize result in "tunnel vision" when examining the professional literature.

A few brief comments are in order before our panel members make individual presentations regarding research in home economics of value to you. More than 75 years ago home economics as a field of study was defined as "in its most comprehensive sense, it is the study of the laws, conditions, principles, and ideals which are concerned on the one hand with man's immediate physical environment and on the other hand with his nature as a social being, and is the study of especially the relations between those two factors." In essence, what research home economists have been primarily concerned with through the years, and more particularly in the last decade, are those research efforts in areas which hold the greatest promise for improving the quality of living for families.

Home economists are concerned with family welfare and home economics researchers have drawn upon their own particular competencies as well as those of many people in allied fields. Our efforts are frequently directed at facets of larger research problems requiring a multi-disciplinary approach. The research that home economists have conducted that you are likely most familiar with, and this is a generalization, is that related to consumer economics, consumption economics, and consumer behavior. We are glad you are reading and critically analyzing research being conducted by home economists in these specialized areas of our broad discipline. However, we also give a good damn about affording you the opportunity to study our research in the more traditional areas of food, nutrition, and clothing as well since

many things are and should be of vital interest to you.

Through a variety of sources of funds, both government and private, home economists are involved in research related to household management, housing, clothing and textiles, equipment, food, nutrition, agricultural economics, chemistry, economic and sociological analysis of rural as well as non-rural living, analyses of public policy alternatives, consumer choice and decision making, and research related to self-concept and social acceptance.

We are beyond looking only at the household in many research efforts and in fact examine many situations in institutional settings. In short, we are doing a lot of interesting and valuable work regarding the well being of families in the United States as well as abroad. The challenging task of our panelists today is to identify several pieces of research conducted by home economists that might be of considerable interest to you even though you may not have examined that work heretofore. We want to convince you of this point: "Home economics research is good, and that's why we should give a damn;"

\* \* \*

Textiles and Clothing

by Patricia J. Sailor

Discussions leading to a Home Economics Research Assessment Planning Projections document began early in 1976. A workshop was held in 1977 in Washington, D.C., to focus on specific problems rather than discussion and assessment on a broad basis. Participating in the workshop were University researchers, University Administrators, representatives from Government agencies including the Social Security Administration, the Department of Labor, the Executive Office of the President, the National Bureau of Standards, the U.S. Department of Agriculture, Department of HEW, and interested consumers. This was a concerted effort by many diverse groups and interests to define problems and suggest areas and priorities for research. Following the HERAPP workshop, the high priority researchable problem areas from each of the research areas were randomized and sent to all workshop participants, most Home Economics Administrators, and to a sample of 4,000 persons selected randomly from American Home Economics Association membership. Recipients were requested to list the ten most important researchable problems using the following criteria: 1) How crucial is the need for a solution to the problem? 2) Do home economists have the expertise to contribute to the solution? 3) Will the solution produce important and widespread benefits? Five major work areas were determined: Human Development, Food and Human Nutrition, Individual and Family Environments, Consumer Economics and Family Resources, and Community Services.

Many projects are currently underway that fall into areas defined as high priority by the study. One area of major concern in Textiles and Clothing was to determine aspects of textiles which affect consumer safety.

As part of a larger study, consumer attitudes toward care labeling and flammability standards were tallied. Results revealed that homemakers want and use care information with 9.5% supporting such labeling. Although 3/4 of the respondents felt flammability standards were needed for

carpets, mattresses, and children's sleepwear, most were not aware of current standards and did not feel a responsibility to take special care to maintain Fire Resistant properties.

Another study concerned itself with the effect of mold on the flame resistance of flame retardant cotton fabrics. The researchers also studied whether mold stains could be removed by conventional methods and what the effects would be on the fabrics. They found that restoration of fabric appearance was not only difficult but may ruin the effectiveness of the flame retardant finish.

Other research with high priority has been for the handicapped. A nursing home served as the locale for research on the clothing needs of elderly handicapped women. It pointed up specific design features that are helpful to such women.

The effectiveness of the 1972 FTC trade regulation concerning "care labeling of textile wearing apparel" continues to be the object of research. One recent study investigated consumer interpretation of care instructions. For the project the definitions of the American Society for Testing Materials were used as a standard to measure the consumer's interpretations of labeling terms. Two of the terms, "wash separately" and "hand wash" were taken more literally by consumers than the ASTM interpretation.

A Home Economics Professor, and Home Economics instructor, a State Cooperative Extension Service Home Economist, and a department store buyer collaborated on a project to determine what sort of dress and appearance creates a favorable impression on potential employers. They found that all the executive secretaries and most of the hiring agents interviewed said that appearance was used as an indicator of the applicant's general ability to meet job requirements. From the comments of these people the researchers were able to isolate some general standards of dress considered appropriate for office wear.

A recent article in the Journal of Home Economics does a good job of describing the type of forward looking approach future Home Economics research will take relating to energy.

Energy-related decisions are present in just about everything we do. In housing, we make decisions on type, size, location, and insulation. We decide where to travel, in what kinds of vehicles, and whether to use public transportation. We make decisions about cleanliness - how long to wear clothes, when to throw the towels in the laundry, and how much hot water to use.

Home economists have a golden opportunity to play a critical role in helping families to make decisions - to look at options and assess both long- and short-run costs of adopting energy-saving habits.

There are no simple answers to the energy problem as one crucial research area, but we can't let that deter us from looking for solutions.

Some useful references for the areas noted here are:

Hogan, M. Janice, "Changing Our Energy Behavior"; Journal of Home Economics, May 1978.

Kelley, Eleanor; Jones, Susan; Nelson, Rogene, "How to Help Your Students Be Successful at Job Hunting"; Journal of Home Economics, November 1976.

Finley, Etta Lucille; et al; "Some Facts About Methyl Parathion Contamination"; Agricultural Experiment Station, Circular #104.

Badenhap, Suzanne B., and Purchase, Mary E., "Laundering Practices and Results of Homemakers Using Coin Operated Laundries"; Home Economics Research Journal, December 1976.

Ahrens, V. D., "Flammability of Children's Sleepwear: Evaluation of Selected Construction Features"; Home Economics Research Journal, June 1977.

Manikowske, Linda and Janecek, Coila, "Low Temperature Home Laundry, Effectiveness and Energy Consumption"; Proceedings of the National Meeting of the Association of College Professors of Textiles and Clothing; October 1977.

Reeves, Wilson A. and Marquette, Yvonne, "Use of Fyrol 76 and Nitrogenous Resins for the Production of Ignition Resistant and Durable Press Fabrics"; Journal of Fire Retardant Chemistry, May 1978

Treece, Anna Jean, and Bore, Joan, "The Effect of Stain Removal Treatments on F. R. Cotton Fabrics Exposed to Cheatomium Globasum, a Standard Mold"; Proceedings of the Association of College Professors of Textiles and Clothing, Regional Meetings, 1976.

Phipps, Cynthia Allen, "Clothing Design for Handicapped Elderly Women"; Journal of Home Economics, September 1977.

\* \* \*

#### The Perspective of Consumer Economics

by E. Scott Maynes

This paper has two objectives: (1) to explain the viewpoint of consumer economics, and (2) to introduce its research domains and questions. It is of necessity brief; those wishing the detailed notes and references should write to the author at the College of Human Ecology, Cornell University, Ithaca, New York 14853.

Though our college has changed its name from Home Economics to "Human Ecology," reorganized its departments, and altered its mix of faculty recruitments, it remains faithful to the Home Economics tradition. This in turn implies a twofold mission:

1. At the micro level: the development, refinement, and transmission of a body of knowledge that enables households/families/consumers to function more effectively;
2. At the macro level: the development and assessment of policies that affect households.

The execution of this mission in the economic domain is the objective of "consumer economics."

The needs of the ultimate client shape any discipline. Let us see how the needs of our ultimate clients--the seller and the consumer--respectively shape marketing and consumer economics.

Both disciplines are applied. For both, theory is an instrument that is justified only by its ultimate usefulness to the client.

Both are multi-disciplinary and eclectic. Because consumer behavior and consumer problems transcend disciplinary boundaries, so too does the composition of faculty and the character of research in both fields. Not to be eclectic might be to court error and result in disservice rather than service to the ultimate client.

Both marketing consumer behaviorists and consumer economists have a common interest in understanding and predic-

ting consumer behavior. But our different clients lead us to different problems or to different views of the same problems. Two examples make the point.

In measuring quality (utility), your concern in conjoint analysis is to ascertain preferences, ex ante and perhaps roughly, with an eye to product planning and selling. We seek to measure quality, ex post with precision, so that we can assess the functioning of markets.

For you the psychology of advertising is a primary tool to be utilized for more effective selling. For consumer economists it is a topic of secondary interest. If pursued, it would represent an exercise in "innoculation": teach consumers the psychology of advertising so that they are less likely to be taken in.

The difference in ultimate client also determines the dominant disciplinary background of our fields. Your primary mission as marketers is to ascertain and influence preferences--a psychological problem. So it is hardly surprising that most marketing consumer behaviorists received their primary training in social psychology.

In consumer economics our primary concern is to find how well consumers are using their resources and how well market and non-market mechanisms are serving them--an economic problem. Hence, training in economics is dominant in our department and, in my judgment, should be dominant among all home economists dealing with consumer problems.

Consumer economics is a developing discipline, just now moving in many areas from description to analysis and characterized by the juxtaposition of primitive and sophisticated research. Its main publication outlets are the Journals of Consumer Research (a small share), Consumer Affairs, Consumer Policy (Europe), Consumer Studies and Home Economics (United Kingdom), Home Economics Research. We deal now with each of four research domains, citing representative topics and authors; not all authors are consumer economists, though this sector of their work is viewed by consumer economics as part of its literature.

Representative topics include values (Carol Meeks and Ruth Deacon), both the theory (Gary Becker) and measurement of household time use (M. Rowe--1917, Jean Warren, Kathryn Walker), financial resource management (credit--Joan Taber and Jean Bowers), other assets (Lawrence Shepard), decision-making within the household (Elizabeth Wolgast, Key Edwards), the long-run saving decision (Franco Modigliani and Richard Brumberg, Martin Feldstein), and use of tangible capital (Keith Bryant).

#### The "New Home Economics"

This phrase, introduced by Marc Nerlove in 1974, refers to the unifying concept of human capital -- its development, use and deterioration. Specific topics include the concept of human capital (Theodore Schultz, Gary Becker), the economics of labor force participation (a topic shared with Labor Economics), of recreation, of discrimination (Barbara Zoloth), of marriage, family formation, and family size.

#### The "New Industrial Organization"

Traditionally "industrial organization" has denoted the efficiency with which industries and markets are organized for purposes of production and distribution, especially the "monopoly problem." New developments extend these concerns to how informationally perfect markets are. Specific topics: the price-quality relationship (Ruby Morris and George Sproles), the measurement of product quality (E. Scott Maynes), the assessment of informationally imperfect markets (Maynes), studies of

actual and optimal search/bargaining (Sproles and Loren Geistfeld and Suzanne Badenhop), the theory of informationally imperfect markets (Peter Diamond, Michael Rothschild, Steven Salop).

#### Consumer Policy

Policies that affect the ultimate client are ever a concern to professionals in any field. So, too, with consumer economics. Worth citing here are attempts to "organize" the field or its questions (Shepard, Maynes, S. J. Ritchie) and examples of research dealing with specific policies, e.g., warranties (Bryand and Jennifer Gerner), consumer representation (Eric Midwinter), a proposal for a local consumer information system (Maynes, James Morgan, Weston Vivian, and Greg Duncan).

#### Conclusion

Even if this menu of research and research topics does not command your interest, the fact that the consumer interest is almost universal makes it imperative that you pay attention to what consumer economists teach and what they research. Potentially at least, consumer economists command the largest audience of all!

\* \* \*

#### Housing Research in Home Economics

by Suzanne Lindamood and Sherman Hanna  
Kansas State University<sup>1</sup>

There is no separate discipline of housing research in home economics. As with many other applied fields, there are housing researchers who teach in schools of home economics, and there are housing researchers who teach in other areas. These researchers use various root disciplines to investigate a wide variety of housing problems ranging from efficient kitchen planning to equitable federal housing policy. The only common thread in all of the housing research in home economics is the paramount concern about families and consumers, as opposed to research conducted from the point of view of planning agencies or businesses.

There are a number of journals which publish housing research articles. Some researchers working out of home economics have published in sociology journals. The primary outlets for home economics housing research are the Home Economics Research Journal and Housing and Society: The Journal of the American Association of Housing Educators. Housing and Society also publishes work by architects, sociologists, economists, and others interested in housing.

Below is a selected list of housing research and publications by home economists. The titles should give some idea of the range of topics covered by housing researchers in home economics.

- I. Southern Regional Research Project (S95) Survey of about 4,000 households in Alabama, Florida, Georgia, Maryland, Oklahoma, Texas and Virginia. Research areas include development of housing quality indexes, environmental factors influencing housing, energy, community and neighborhood characteristics and preferences, subjective and objective measures of quality and satisfaction, housing needs of the aged, housing goals and aspirations, subcultural differences, constraints to obtaining adequate housing, single person households, and desire for help with housing problems.
- II. North Central Regional Research Project (NC128) Quality of Life as Influenced by Area of Residence (metro, metro fringe, non-metro). Research areas include how people manage their housing problems,

norms and deficits, and assessment of conditions and needs.

<sup>1</sup>The authors are members of the Department of Family Economics, Kansas State University

III. Representative articles in Housing and Society (formerly Housing Educators Journal).

Energy Policies Directed at the Home: Which Ones Will People Accept?

Space Norms and Housing of Low Income Families

Housing and Occupational Subcultures

Social Factors Related to Housing Selection

Attitudes of Blacks and Whites in Public Housing

The Measurement of Quality in Housing and Its Relation to Housing Satisfaction

Housing Choice and Distance Moved: An Ecological Model

Simulations of Variable Rate Mortgages

The Effects of Housing Allowances

The Social Psychology of Space: Measuring Territorial Behavior of Elderly People in Public Housing

Residential Homogeneity and Neighborhood Satisfaction

Residential Satisfaction of Recent Movers into Government Assisted

Housing Projects: The Impact of the First Nine Months

IV. Textbooks

Morris, Earl W. and Winter, Mary. Housing, Family and Society, Wiley, 1978.

Lindamood, Suzanne, and Hanna, Sherman. Housing, Society, and Consumers: An Introduction, West, 1979.

V. Selected Articles in Home Economics Research Journal

Consumer Preferences and Selected Socioeconomic Variables Related to Physical Adequacy of Housing

Housing Values, Aspirations, and Satisfaction as Indicators of Housing Need

Service Life of Appliances: Variations by Selected Characteristics of Owner Households

Home Maintenance and Improvement Behavior of Owners

Housing Decisions in Selecting a Residence in a Planned Townhouse Development

Housing Aspirations of Southern Appalachian Families

VI. Selected Articles in The Journal of Consumer Affairs

Evaluation of Energy Saving Investments

Economics of the Low-Income Mortgagor Counseling Program

VII. Selected Dissertations

Socio-Physical Factors Affecting Energy Consumption in Single Family

Dwellings: An Empirical Test of a Human Ecosystems Model



OVERVIEW OF THE "LONGITUDINAL RESEARCH DESIGN  
AND ANALYSIS OF PANEL DATA" WORKSHOP

Dr. Fred Cutler, Kennedy Research, Inc.

Introduction

This workshop was organized with two primary goals in mind. First, there is a need to advance consumer behavior theory through the implementation of a longitudinal research designs. At present, consumer research primarily utilizes a "snapshot" approach which examines the environment at only one particular point in time. There are alternative research methodologies available akin to a motion picture running across time; dynamic rather than static in nature. Longitudinal research designs allow for a response gathering situation in which you are reinterviewing the respondent over time, and/or acquiring response data over various time sequences.

The second purpose of this workshop was to present some methodological problems that are associated with longitudinal designs. It was felt that there is insufficient attention being paid in the literature toward the discovery of plausible solutions for some of these problems and that pointed discussion might suggest fruitful avenues to pursue. As a start, problems associated with several studies now in progress were identified, and potential solutions to these problems were advanced.

This workshop was also an example of the growing importance of the subject of "time" in consumer behavior theory. There are going to be at least two different perspectives that researchers will be using to look at the importance of time as it relates to consumer behavior theory. One is the consumer's time perspective. For example, Jon Gutman and Don Kanter at the University of Southern California have been looking at what is called planning horizons, or an individual's outlook towards the future when he is about to make a purchase decision. A second perspective relates to measuring change in attitudes and/or behavior over time. This workshop looked at the issues of longitudinal research designs which are able to measure attitudinal and behavioral change over time.

The Participants

The panelists for this workshop discussion were:

- Dr. Lois Benedetti, AT&T Long Lines
- Dr. Fred Cutler, Kennedy Research
- Mr. Randall Harris, National Family Opinion, Inc.
- Dr. Ed Hart, Heublein, Inc.
- Dr. Don Lehmann, Columbia University
- Dr. Lorraine Scarpa, Heublein, Inc.
- Dr. Arch Woodside, University of South Carolina

Discussion in the workshop was wide-ranging, and difficult to easily summarize. What follows is a brief description of some of the issues which were raised, and which might prove to be provocative for some interested readers. In addition, a brief reference listing is provided for those who wish to delve more deeply into this area.

Multi-Wave Measures

There are many stated advantages to longitudinal data. One of the most basic is the possibility of measuring cause and effect. If you measure over time, you can

theoretically see which comes first, the chicken or the egg. While there are recognizable difficulties in achieving this elusive goal, you are certainly better off with longitudinal data than you are with a simple snap-shot.

Longitudinal designs also allow us to start to monitor decision processes over time. This is important for researchers interested in why people make a choice decision -- not just the way they make it in a snapshot time period, but the ways decision processes change over time. One result of static research seems to be that the notion of the way consumers choose somehow gets simplified. In opposition to many of the prevailing implicit views of much consumer decision-making research, it should be understood that people do not just sit and process massive amounts of information each time they make a purchase.

There are several consumer variables that one can track via longitudinal designs. The most common procedure for tracking variables is through mean levels, such as average purchase rate or average brand share over time. A related statistic is the track of the variance or the standard deviation of some measure of how well people agree on the ratings of the brands, etc. In general, variances will decrease as a new brand gets learned about.

There are, however, problems in simplistic applications of summary statistics. If, for example, you measurement is on a 1 to 10 scale and people all start out at a 1 on the scale (meaning, "I have never heard of it," or "I don't know anything about it"), the variance has to increase simply because of the constraints of the scale.

Similarly, another interesting aspect of multi-wave data is that the non-response percentage on a question will drop considerably as something becomes better known. In a recent five wave study, the telephone panel was questioned on a new brand and an old brand. Non response tended to be around 4% for the existing brand. For the new brand, it started out around 25%-30%. However, after four or five waves of interviewing, which took about 18 months and coincided with the introductory period of the product (Vega), the item non response percent dropped to the 4% level. One can notice a brand getting learned just by the percent "don't know", "didn't response," "refused to answer the question," etc.

Another interesting construct that researchers can follow is change in consumers' attitude structures. Assuming a new brand, the first step toward uncovering structures might well involve factor analysis. Generally, only one factor will surface. Respondents either tend to like it or they don't. As time goes on, the single factor starts to break into two or three factors, reflecting that the way consumers approach brands seems to change over time. This is something you cannot easily spot in cross-sectional data, but is readily apparent through longitudinal data.

Yet another exercise that is of interest is to track

model parameters. For example, if one tries to build a model of how awareness relates to comprehension and comprehension to attitude and attitude to choice, a hierarchy of effects model is often utilized. Analyses here relate to the stability of the parameters across groups and across time. Longitudinal data allows one to examine this type of issue, and many others as well.

#### Problems With Longitudinal Designs

There are a number of problems in this area, some well-recognized and others more hidden. For example, people can become conditioned by being asked questions, and using the same people over time can become a problem. Also, not everyone agrees to participate over time, suggesting a potential "dropout bias." What happened to the people that dropped out, are they different or not different?

Still another problem relates to what could be labeled "adjustment time." Consider adjustment time within a hierarchy of effects model, where one believes the effect of awareness on attitude is realized within a day or less. If you are measuring every six weeks, it becomes difficult to search for cause and effect looking at data the way we traditionally look at experimental data. If, on the other hand, it takes six or eight weeks after you become aware before you comprehend something, then the longitudinal designs can work a lot better. The last problem is one of collinearity. For any kind of semi-stable market, the correlation between time one and time two on most measures, is going to be high. Trying to "pull out" what's going on then requires some fairly sophisticated estimation procedures.

#### Problems With Consumer Panels

The theoretical literature on longitudinal consumer panels has concentrated on several perceived problems with this methodological approach. The one problem that researchers talk about most is their representativeness of panel data. However, some recent studies have shown that the demographic and psychographic (especially the psychographic) characteristics of people on a panel are no different from the characteristics of people who do not join a panel. Very little difference has been found in studies conducted by National Family Opinion (NFO). These results contradict much of the academic literature.

Related to this is a concern about the use of "professional" panelists. Studies that NFO has recently conducted using matched samples of people that have been on a panel for one year versus people that have been on a panel for five or ten years, found the level of purchasing to be no different. NFO has looked at the people that drop-off a panel and found that they were not different demographically.

Do you make panel members more aware of what they are buying -- do you make them price conscious, for example -- by asking them to respond over several occasions? These potential problems are a legitimate concern, especially with such studies as a gasoline diary where people record price per gallon. However, it was also argued that use of a one-time or cross-sectional study creates more problems than using people who have been on a panel for a long period of time.

The most severe problem with longitudinal panel studies however, involves "memory pegs." Studies have shown significant differences between diary data and a one-shot multi-card or caravan study based upon a single recall question (i.e., respondents are asked what they have purchased in the last month).

NFO throws out the first two diaries that are returning by a new panel member. These initial "learning" diaries are used to enforce that respondents understand how to fill out the diary. Response overstatement is often seen in these first two monthly periods. Much to the chagrin of the client who says "I'm paying for that, why can't I have the data" - the researcher is saying that data may be lacking in quality. Respondents are not used to filling out the diary and therefore the data is questionable.

As suggested above, however, it has also been found that when one conducts one-shot research, there is often an overstatement of purchase behavior. In many instances, when people are asked what is the major brand of gasoline they buy or what is the brand they buy most often the major brands (i.e., those that advertise) come out twice as high as their market shares would suggest.

#### Use of Control Groups

All longitudinal designs have the inherent problem of test bias. Respondents are subjected to numerous interviewing situations, each of which can and often does bias the results of the succeeding interview. The use of a control group that is not subjected to questionnaire stimuli becomes almost mandatory to allow for the measurement of possible test bias.

Numberous advantages can be cited for a long period for a longitudinal design. However, if a control group is administered some form of a pre-test and then re-interviewed via a post-test questionnaire, this long time period produces operational problems. Such a long time period makes it very difficult to locate control respondents who have since moved or changed telephone numbers. Additionally, it would often be desirable to include a second control group that is only the advantage of measuring the effect of pre-test administration with the primary control group.

#### Conclusions

Longitudinal research designs offer unique advantages over cross-sectional designs for consumer research studies. Longitudinal designs appear uniquely suited for those studies that are attempting to uncover changes in consumer attitudes or behavior across different consumption situations and/or merely across time.

However, the problems associated with the proper administration and eventual data analysis of longitudinal studies are many. Some of these problems were discussed in the workshop, with positive suggestions for resolving many of the administration and analysis difficulties.

#### References

- Campbell, Donald T. and Julian C. Stanley. Experimental and Quasi-Experimental Designs for Research, Chicago: Rand McNally, 1963.
- Enis, Ben M. and Gordon W. Paul. "Psychological and Socio-Economic Atypicality of Consumer Panel Members," presented at the Fall Educator's Conference, American Marketing Association, August 26, 1969.
- McCullough, B. Claire. "Effects of Variables Using Panel Data: A Review of Techniques," Public Opinion Quarterly, February 1978, pp. 199-219.

Neyer, John and Joseph Waksberg. "Conditioning Effects from Repeated Household Interviews." Journal of Marketing, Vol. 28, April 1964, pp. 51-56.

Parfitt, John. "How Accurately Can Product Purchasing Behaviour Be Measured by Recall at Single Interview," presented at XX Esomar Conference, Vienna, Austria, August 1967.

Powers, Edward A., Willis J. Goudy and Pat M. Keith. "Congruence Between Panel and Recall Data in Longitudinal Research," Public Opinion Quarterly, June 1978, pp. 380-389.

Pratt, Robert W., Jr. "Understanding the Decision Process for Consumer Durable Goods: An Example of the Longitudinal Approach," Marketing and Economic Development, ed. Peter D. Bennett, American Marketing Association, 1965.

OVERVIEW OF: "TIME: THE FUNDAMENTAL THINGS APPLY"  
(SPECIAL TOPIC SESSION)

Rebecca H. Holman, Pennsylvania State University  
M. Venkatesan, University of Oregon

Introduction

This special topic session consisted of a panel of six individuals doing research on the topic of time and consumer behavior. They are: Douglass K. Hawes, University of Wyoming, Phillip Hendrix, Florida State University, Rebecca H. Holman, Pennsylvania State University, Robert B. Settle, San Diego State University, Christine Urban, Harvard University, and M. Venkatesan, University of Oregon. The session was a direct outgrowth of informal meetings among some members of the panel at the 1977 Association for Consumer Research Conference. While there is increasing recognition that temporal dimensions of consumer behavior should be studied systematically (Jacoby et al, 1976), no major conceptual treatment has so far been focused on the topic. It also became clear that some of the research that is presently underway by some members of this panel related to "time perception," while some of the others were engaged in studying the allocation of time by consumers, often called "time-budget studies." The organizers of this session wanted to merge these two streams of emerging research in this area and to expose the membership of ACR to the variety of perspectives taken in the study of time as a variable in consumer behavior research. It was felt that such an exposure would be more meaningful than presenting an inventory of research findings.

The session was organized so that more general, theoretical issues would be dealt with first, proceeding through specific applications, to overviews of broad areas of research. The subjective-objective dichotomy was present throughout the session, however, with the panel showing roughly equal attention to these two dimensions. The balance of the paper summarizes the remarks made by members of the panel and is arranged (not chronologically) in such a way to take the reader first thru a review of the variety of treatments of time as a variable, some perspectives on perception of time and individual differences associated with it, then presenting illustrative views of "objective" approaches currently being utilized and ending with a critique of "time budget" studies and some of the problems associated with conducting research in this area.

Overview of Time and Consumer Behavior

Douglass Hawes' remarks provided needed perspective on the topic. He reviewed the large number of ways in which time has been conceptualized and how time is seen as a variable in various models of consumer behavior, suggesting directions for future research. (Also see Hawes 1978 for more elaboration.)

Seven different meanings of time were identified by Hawes. These are as following:

- a. Baseline: Time is a way in which activities are evaluated. Pleasant activities pass quickly, unpleasant ones seem to take an inordinate length of time.
- b. Benchmark: Time provides a means for recollecting or anticipating the sequence of activities relative to one another.

- c. Reference point: By knowing the "objective" time one knows what one ought to be doing and where that activity should be done; time locates the individual in space.
- d. Biological clock: The organism maintains an internal rhythm which can be conceived of as a physiological clock.
- e. Drug-altered: Through the influence of certain drugs, one's subjective experience of time can be altered.
- f. Constraint on activities: Time limits one's activities, especially activities that may be discretionary (leisure activities).
- g. Resource: Much like money and other resources, time is something to be "saved" "spent" "lost" or "sold," and different amounts of this resource may be an indicator of one's relative affluence (or poverty).

In reviewing nine major models of consumer behavior, Hawes found a variety of treatments of time as a variable, although all models treated time in some way. (See Hawes (1978) for extensive discussion.)

Areas needing future research were suggested by Hawes. The bulk of the research to date has been done on time budgets and Hawes called for a movement beyond time budget research perhaps looking at how some of the perceptual or subjective dimensions affect time budgeting activities. A merger of the two fundamental streams of research represented by the panel members is clearly indicated.

In suggesting specific topics of interest, Hawes mentioned cultural determinants of time orientation citing Whorf's work (1966) with the Hopi as indicative of the effort of culture upon both perceptual processes and budgeting activities. Hawes also mentioned the individual's personal value of time and how this impacts upon the allocation process (activities allotted a measure of time as they are variously valued).

Subjective Experience of Time

Robert Settle and his colleagues (Pamela L. Alreck, Michael A. Belch, John W. Glasheen, and Robert W. Haas) are interested primarily in individual differences in time orientation and how these individual differences are related to a variety of consumption behaviors. Alreck (1976) developed the F-A-S-T Time Orientation Test and in the session, Settle explained the dimensions tapped by this instrument. F-A-S-T stands for the four dimensions that are the focus of this test, viz, focus, activity, structure, and tenacity.

Focus refers to an individual's tendency to experience events along the full spectrum of a time continuum. Some individuals tend to spend more time remembering (and reliving) past events while others imagine (and live) a not-yet-experienced future. The remainder of individuals direct their consciousness at present events

instead of those in both the future and the past. In scoring this and all other dimensions of the F-A-S-T, individuals instead of being typed are arrayed along a continuum, the end points, for Focus being future and past orientation.

Settle pointed out in later discussion that the Focus scale is not necessarily age-dependent. Intuitively one might think that older-aged individuals, with a limited real future, would necessarily turn to the past for images. This is not necessarily the case, as an improbably imaginary future can be as compelling as one that has a high probability of being realized.

Activity refers to the individual's perception of the supply of time available for use. As one end point of the Activity spectrum is the perception that time passes slowly, and that there are not enough activities to fill into the unit of time available.

Structure concerns how an individual imposes (or fails to impose) rigid units upon the time dimension. Highly structured individuals are those who conceptualize time in discrete units anchoring events in terms of objective measurements of time (hours, days, years, etc.). Highly structured individuals have specific "times for" certain activities (meal time, work time, leisure time, bed time, etc.), whereas those low in structure tend to engage in the same activities whenever it seems appropriate, not at fixed intervals. Individuals low in structure see time as a commodity without boundaries and engage in activities accordingly.

Tenacity refers to the degree to which an individual has a need for extrinsic rewards at short intervals of time or whether longer intervals between rewards will suffice. Individuals able to "delay gratification" consider time as irrelevant when engaging in activities, whereas those needing a frequent and continuous schedule of reinforcements in order to sustain an activity consider the length of time between reinforcements critical.

Settle pointed out that the scale has been validated with a large number of adult subjects. There is also high reliability with the scale using a number of different measures of reliability. Additionally, the scale has been used to differentiate life styles among consumers (Settle, Alreck and Glasheen 1978).

In contrast with Settle, et al, Rebecca Holman (and her colleagues, Roger D. Evered, The University of Illinois and Michael D. Reilly, The Pennsylvania State University) are more concerned with future orientation. Since the future is a purely imaginary event, it is reasonable to assert that individuals have different and systematic ways of constructing that imaginary future, linked in some way with other cognitive processes. In operationalizing Evered's conceptual view (Evered 1973) Holman et al have developed the Futurizing Style Questionnaire (FSQ). Three types of ways that individuals conceptualize the future were identified by Evered. Holman explained the ideal types of each of the three.

Participants see the future as a logical, linear extension of the present. They have a clear image of how the future will be different from the present and can easily see the steps needed to get from the present to the future. Participants can easily work within the systems in which they find themselves, using those systems as a means for achieving the futures they see for themselves.

By contrast Prospectors can envision many varied futures and can deal easily with a future that might be radically different from the present. They entertain

the notion that the future might be discontinuous with the present and therefore see no necessity for a logical connection between the present and the future. Whereas the participant plans for the future by setting down goals and objectives that move one closer to one's goals, the prospector plans for the future by keeping as many options open as possible and by acquiring a multitude of capabilities to be able to cope with the variety of things that "might" happen.

Producers, on the other hand, see the changes that occur in the future as superficialities. Producers see the future as being fundamentally like the present. To the producer, the only things that change over time are the things that do not matter, and if there is basic change, it is gradual over long periods, much greater than the span of one person's life. Planning for the producer means devising a strategy to cope with the status quo and perfecting that strategy over time.

Holman pointed out that the FSQ has been validated with only a limited population, but that a French translation had been made (by Alain Jolibert, University of Grenoble) and that data from French students was currently under analysis. Holman suggested the linkage of language with ability to imagine a future and suggested this as a further area of potential fruitful activity.

#### Allocation of Time by Consumers

Sociologists have been interested in the amounts of time people spend in different activities. They were interested in all the activities of individuals, which included what we would term as "consumer-behavior related" activities. Such studies have utilized a "time-budget" method, whereby time allocation for the entire day (24 hours) for different activities are ascertained. The well known work in this area is by Szalai (1972). The two major studies that report time allocation in the United States are by Chapin (1974) and Robinson (1977). Robinson's data comes from the Survey Research Center, Institute for Social Research, of the University of Michigan, which has been collecting time allocation information for some time as part of their "Time Use Study."

Hendrix's study is based on the data collected by the Institute for Social Research. However, his presentation at this session and his paper published elsewhere (in its entirety) in this volume of the ACR Proceedings are intended to provide a conceptual framework for analysis of the time budget data. His conceptual approach is based on the view that if we analyze "why people spend various amounts of time, then we have good ideas as to which goods and services relate to various activities." He provides a framework for classifying the activities and then provides a model for allocation of time based on his classification (See paper by Hendrix, Kinnear and Taylor for details).

#### Time Allocation and Media Use Patterns

Robinson and Converse (1972) examined the developmental changes in media use patterns as a function of the increasing intrusion of television. Others (e.g. Chapin, 1974) have considered television watching as a passive form of activity. Christine Urban focused her remarks on the allocation of time for media usage -- a form of consumer behavior that is not only important in defining an individual's life-style and allocation of "disposable" time for other activities, but is also, in itself, a most time-intensive and time-ordering activity.

She pointed out that one can examine volume of media use as a means of gaining understanding of how different people 'order' their day. Recent studies have found

that there are sub-groups within the media 'market' that are obviously 'heavy' or 'light' consumers of media. In addition, these findings seem to suggest that: (1) The volume of media consumption not only affects "real-time" allocation for other activities, but also perceptions of available (disposable) time; (2) That heavy media involvement affects (or is effected by) a greater stability in daily routine than is evidenced in the group as a whole; and (3) That the definition of each medium (e.g. as an "active" or "passive" way of spending time) differs for these heavy and light groups. "Media-heaviness," then, might provide a useful explanatory variable for research investigating time use and orientations: providing clues about the priorities, motivations and characteristics that distinguish different "time-user" segments.

Secondly, since media play such an important role for social integration, much of the 'purchase decision' (especially for television) appears negotiated according to the "least-objectionable" rule (i.e. no one in the decision-making unit usually maximizes satisfaction), we can hypothesize that negotiated time-use decisions will provide more stable patterns, and will be less reflective of an individual's true preferences. This distinction has important consequences when analyzing time-budget data.

A third area of findings in media research is that of cyclicity of overall preference trends in the total audience, and the direct interdependence of media trends and time-use trends. Just as certain modes of presentation and topics have a natural "life-cycle" within the media, so also are there fads or fashions in time-use (e.g. jogging, handicrafts, etc.). These are strongly interrelated, and it could be profitable to understand the influence that media coverage has on the adoption cycle of certain behaviors or activities - especially leisure time activities.

Looking at these kinds of findings within media research, then, can provide not only insights into the rich, multi-dimensional behavior toward a most time-sensitive activity media use, but also highlight more general factors to be investigated in research on time-budgeting and orientation.

#### Critique of Time Budget Studies

Venkatesan pointed out some of the major difficulties associated with the collection of time allocation data for consumers. He said that while Szalai (1972), Chapin (1974) and Robinson (1977) had collected time allocation information for all of the activities, our interest is in obtaining information related only to consumer behavior. While these "time budget" studies had obtained information on all sorts of activities, they invariably collapsed information relevant to consumer behavior into one category and named it "marketing activities."

The second problem faced by researchers attempting to get time budget information is the "classification" of activities. Activities have been grouped into either "primary" and "secondary" or "obligatory" and "discretionary." Many activities of interest to consumer behavior researchers occur simultaneously, such as talking about a product or consumption activity while watching television, and the like. The binary classification system is inadequate. Even the three category classification proposed by Hendrix (1979) does not fully solve this problem of getting time allocation information for activities occurring simultaneously. Another approach used by some has attempted to get time allocation information from respondents based on the chronological order of all of the activities during a 24-hour period. This approach, too, may miss information on

concurrent activities. Thus, a better classification is needed for research in this area that may be useful for consumer researchers.

Venkatesan next addressed his comments on the method utilized for data collection. Much of the time-use data is obtained from respondents for a 24-hour duration relating to all the activities taking place during this time period. Other techniques have utilized a "yesterday" format, in which respondents are to recall all the activities and their corresponding time allocation; or "tomorrow's activities" format, in which case a "diary" is left with the respondent with instructions and which is collected the day after its completion. Different proportions of the sample are contacted for differing days of the week, thus assuring data collection for the whole week. While the data may be aggregated for the whole week, such information is not likely to reveal any patterns or variations in consumer behavior. Such patterns can only be discerned from information collected from the respondents for a longer period of time and preferably for different periods of the year. Such a requirement points to the use of consumer panel methods to obtain consumer time-use information. However, he pointed out that panel method introduces its own peculiar problems and the costs associated with setting up a panel and its maintenance would preclude many researchers from collecting time-use information.

Another problem relating to the collection of time-use information, according to Venkatesan, is that all the information comes from "self-reports" from consumers. Thus, they may contain data of varying reliability and validity. Other measures must be utilized as "reliability checks." There are other problems of data collection in this area, such as "units of recall," inclusion of travel time, and appropriate ways of including "window shopping" and other time-use activities. Venkatesan concluded by saying that these problems that are indicated are not insurmountable, and in fact this area appears to be most promising for research, as the data obtained on time allocation is amenable for varying levels and elegance of analysis.

#### Summary and Conclusions

All participants in the session (both panel and audience) seemed to find the session valuable and stimulating. Panel members invited audience members to contact them if mutual interest was uncovered. Offers were made to trade papers and references, the latter being elaborated upon by Robert Settle who offered to coordinate an annotated bibliographic exchange. Rebecca Holman offered to look into the possibility of a special mini-conference on the topic if sufficient interest was generated in the future. In general, it was felt that the interaction among those interested in the topic of time was quite valuable and that it was rewarding to know that others were engaged in complementary research.

#### References

- Alreck, Pamela L. Time Orientation and Behavior. Unpublished manuscript, San Diego State University, 1976.
- Chapin, Jr., F. S. Human Activity Patterns in the City. New York: Wiley, 1974.
- Evered, Roger D. "Conceptualizing the Future: Implications for Strategic Management in a Turbulent Environment." Unpublished doctoral dissertation, University of California, Los Angeles, 1973.
- Hawes, Douglass K. "The Role of Time in Models of Consumer Behavior," Research Paper No. 270, Institute for Policy Research, College of Commerce and Business Administration, University of Wyoming, October, 1978.

Hendrix, Philip E., Thomas C. Kinnear, and James N. Taylor. "The Allocation of Time by Consumers" in the Proceedings of the Association for Consumer Research, W. Wilkie (editor), 1979.

Jacoby, J., G. J. Szybillo and C. K. Berning. "Time and Consumer Behavior: An Interdisciplinary Overview." Journal of Consumer Research, 2-1 (1976), 320-339.

Robinson, J. P. and P. E. Converse. "The Impact of Television on Mass Media Usage: A Cross-national Comparison" in A. Szalai (editor) The Use of Time. Paris: Mouton, 1972.

Robinson, J. P. How Americans Use Time. New York: Praeger Publishers, 1977.

Szalai, A. The Use of Time. Paris: Mouton and Co., 1972.

Settle, Robert B., Pamela L. Alreck and John W. Glasheen. "Individual Time Orientation and Consumer Life Style." H. Keith Hunt (editor) Advances in Consumer Research, V. Ann Arbor: Association for Consumer Research, 1978, 315-319.

Whorf, Benjamin L. Language Thought and Reality. Cambridge: The M.I.T. Press, 1956.

# THE ALLOCATION OF TIME BY CONSUMERS<sup>1</sup>

Philip E. Hendrix (student), University of Michigan  
Thomas C. Kinnear, University of Michigan  
James R. Taylor, University of Michigan

## Abstract

A model which parallels the process by which consumers allocate their time is proposed. Essential concepts, such as the relationship between goods and services and consumers' time, the classification of activities, etc., are further refined.

## Introduction

The notion of life-style has been applied extensively in the study of consumer behavior. (Wells and Cosmas, 1977) Though no definition is generally accepted (Ferber and Lee, 1974), the essence of the concept is well conveyed by the following:

Life style research is a sociological concept which deals with time and energy (and money) allocation where the individual has a choice of one activity or another. (Demby, 1971)

The preeminence of the resource of time in the life styles of consumers has been suggested by Nicosia and Mayer (1976, p. 68):

It may well be that time is the crucial dimension [of consumption activities] since its supply is fixed. The interdependence among all kinds of activities because of the fixed supply of time has strong implications for the study of a society's consumption.

And yet, despite the integral nature of time in consumer behavior (Schary, 1970), no major conceptual treatment or systematic empirical effort has emerged. (Jacoby et al, 1976, p. 330)

This paper is intended to partially fill this conceptual void by developing a framework from which subsequent empirical analysis of the allocation of time by consumers may proceed.

## Activities As A Process

The view of activities as productive processes is an integral element of the theory underlying consumers' allocation of time. (Becker, 1965) Essentially, an activity may be regarded as a process associated with which are certain desired outcomes. Also associated with activities are the inputs of time, energy, and goods and services which may be combined in varying proportions to effect the desired outcomes. To illustrate these concepts, the activity of meal cleanup may be utilized. The desired outcome associated with this activity may be the restoration of the dining area, dishes, utensils, etc., to some satisfactory level of cleanliness. This may be achieved in a variety of fashions, each of which represents a different combination of the inputs of time, energy, and goods and services. For instance, one may hold the energy and goods inputs to a minimum at the expense of time by washing the dishes and utensils in the traditional manner at a relatively leisurely pace. Alternatively, one's time and energy inputs may be held

to a minimum by incorporating such goods and services as disposable plates and utensils, a dishwasher, or even a maid. Other activities - meal preparation, procurement of goods, maintenance of one's possessions, recreation, etc. - may be thought of similarly.

From this perspective, one can see how both the types and levels of desired outcomes may affect the mix of inputs. For instance, an individual whose set of desired outcomes includes both physical conditioning and socializing may select an entirely different activity - perhaps tennis - from that of an individual whose only objective is physical conditioning. The mixes of time, energy, and goods and services are likely to differ as well. Similarly, one who engages in various activities - say, meal preparation and maintenance of home and auto - in hopes of contributing to his or her sense of accomplishment may be expected to utilize a mix of inputs unlike those of an individual who has no such objectives. We might expect the latter to spend less time in such activities, perhaps opting for prepared foods and the services of the local handyman and mechanic, respectively, as inputs to the above activities.

One other aspect of this framework is worth amplifying - that is the relationship between goods or services and the time input to activities. Depending upon whether the absence of the good or service would increase or decrease the amount of time devoted to a given activity, goods or services may be classified as either substitutes or complements, respectively, for the time input to that activity. Thus, a telephone may be a substitute for the time spent shopping and a complement to the time devoted to corresponding with distant relatives. Though classification rests upon the resulting impact on time, general patterns of usage suggest that substitutes include such items as partially prepared foods, restaurants, dishwashers, aluminum siding, and automatic car washes, while complements include gourmet cooking classes, tennis rackets, theatres, and televisions, to name a few.

## The Equilibrium Model Of The Allocation Of Time

Having established the notion of activities as productive processes, Becker (1965) suggests that the inputs of goods and time will be combined according to the cost minimization rules of the traditional theory of the firm. Amplifying this proposal, Linder (1970, p. 44) proposed that "time will necessarily be distributed over different sectors of use in such a way that the yield on time is everywhere the same." (1970, p. 44) Thus the optimal allocation of one's time would equate the marginal utility associated with each of the activities in which one may engage.

Because time is apportioned within a fixed budget constraint, the choice of activities to which time is devoted creates opportunity costs. Economists have traditionally valued these opportunity costs at the individual's rate of pay. (Evans, 1972, p. 1)

<sup>1</sup>The authors would like to thank F. Thomas Juster, University of Michigan, for his insightful comments.

In classical theory, opportunity costs [are] the goods and services one does not buy, lacking the wages that could have been earned had the [time]



been spent at work. (Bell, 1975, p. 559)

These opportunity costs are then incorporated into analyses of a broad range of consumer behaviors, such as locus of food consumption (Prochaska and Schrimper, 1973), mode of travel (Devany, 1974), family size (Sawhill, 1977), and even participation in organized religious activities (Azzi, 1975).

Implicit in this approach is an assumption of virtually complete latitude in the allocation of one's time. Such an assumption belies reality for a significant number of individuals and activities. Nor has this assumption escaped criticism. (Morgan, 1968; Mabry, 1970; Hameed, 1972; Evans, 1972) Bell summarizes the argument against the traditional valuation nicely:

One difficulty with the current economic theory of time as the basic unit of expenditure for both consumption and work is that it assumes an easy substitution between hours of work and hours of consumption. Such a state of affairs does not, of course, describe reality for the vast majority of earners, whose opportunities to increase or decrease their hours on the job are sharply circumscribed. The alternative to one kind of consumption activity may not, realistically, be using the time at work but spending it in another form of consumption activity. (1975, p. 563)

Even the choice of non-work activities to which time may be allocated is circumscribed. What may be called anterior conditions compel one to devote some amount of time to certain activities. For instance, physiological requisites stipulate at least some minimum amount of time for sleeping and eating. Similarly, the parent of two small children has certain time commitments, *ceteris paribus*, which a childless individual either does not have, e.g., child care, or has to a lesser degree, e.g., meal preparation or laundry. Similar relationships between anterior conditions and the time input to various activities are depicted in Table 1, where again "mediating factors" (to be dealt with subsequently) are assumed constant.

TABLE 1

HYPOTHESIZED RELATIONSHIPS BETWEEN TIME INPUT TO VARIOUS ACTIVITIES AND SELECTED ANTERIOR CONDITIONS

Activities	Anterior Conditions	Probable Relationship
Personal care	Age Health	positive higher if poor
Household Cleaning	Size of dwelling Age of dwelling Stage in FLC	positive positive curvilinear
Shopping	Stage in FLC Proximity to facilities	curvilinear negative
Meal preparation	Stage in FLC	curvilinear
Meal cleanup	Stage in FLC	curvilinear
Laundry	Stage in FLC	curvilinear
Yard maintenance	Size of yard	positive
Auto maintenance	Age of auto Use of auto	positive positive
Child care	Stage in FLC	curvilinear

Of course, the latitudinal constraints posed by these conditions may dissipate or be alleviated over a longer

period of time - children mature, people change jobs, families move,<sup>2</sup> etc. Other conditions are likely to take their place, however. While the children have grown, the health of the spouse or grandparents may have deteriorated. Or the newer, larger home with the expansive lawn may require even more upkeep than the preceding residence. Thus, even one's non-work time must inevitably be allocated within the latitudinal constraints posed by various anterior conditions.

Because it ignores these constraints, the equilibrium model of time allocation is an inadequate approximation of reality for our purposes. Before specifying a more tenable model of time allocation, however, a taxonomy of activities consistent with this notion of latitude must be developed.

#### Taxonomy Of Activities

A prerequisite to the modelling of time allocation is the classification of activities in some meaningful fashion. This is no easy task. Attempting to classify activities as either obligatory or discretionary, Chapin (1974, pp. 70-71) noted the difficulties:

...An exact assignment of an activity to either category would require interpretation of the activity in the context of a number of contingencies - meaning ascribed to that activity by the culture and the social system in which the subject falls, the culturally defined household role being performed by the subject at the time of the activity, the subject's own taste and preference, situational factors surrounding the occurrence of each such activity, [and the time frame].

Before settling on a somewhat arbitrary method of classification, Chapin experimented with a number of alternative approaches. In one, subjects were asked to classify their own activities. Variability in the definition of the concept "discretionary" employed by subjects made for inconclusive results, however. In another approach, supplemental data were gathered to enable the researcher to classify the activities according to predetermined criteria, such as how long prior to an activity had a decision been made to engage in that activity. Respondent fatigue rendered this approach unacceptable, though. Robinson (1977, p. 134) suggests three attributes - enjoyable, discretionary, and spontaneous - which distinguish leisure or free-time activities from obligatory activities. Citing a University of Michigan Institute for Social Research study which attempted to identify the discretionary and spontaneous elements of activities, Robinson noted:

While there is some correspondence between the discretionary nature of an activity and its spontaneous or enjoyable character, ... it is far from complete. It is hoped that these [indeterminate] results will stimulate further research into operational definitions of [uses] of time. (1977, pp. 137-138)

Regardless of the means by which activities are classified, a dichotomy is entirely too restrictive. Dichotomies of activities, such as labor-leisure or obligatory-discretionary, ignore the set of activities which contain elements of both obligation and discretion. As

<sup>2</sup>Michelson (1973), examining the relationship between moving behavior and deficits in one's socio-physical environment, posits residential moves as a response designed to alleviate these constraints.

DeGrazia (1962, p. 93) has noted, "things with a sense of obligation may also be felt as enjoyable," citing the care of children as an example. Bell (1975, p. 559) lends support by noting certain activities - sleeping, eating, etc. - which cannot be neatly assigned to either category of the "artificial leisure-labor dichotomy." Finally, Chapin (1974, p. 37) himself has observed:

In between obligatory and discretionary, the assignment of an activity to one or the other category can vary with the person and the circumstances.

There is no reason to cling to such a clearly inappropriate taxonomy. Extending the propositions developed earlier, activities arrayed along a continuum according to the degree of latitude one has in determining in the short run the amount of time so employed<sup>3</sup> fall into three somewhat discrete categories. At one end of the continuum are inelastic activities in which the individual has no latitude in determining the time input. Employment, for some, may fall within this category (though not for others, e.g., professionals or self-employed, who may exercise some degree of latitude.) At the other end of the continuum are discretionary activities in which one may exercise virtually complete latitude in so employing one's time, e.g., socializing, recreation, etc. In between these extremes are intermediate activities, with which is associated some latitude. Intermediate activities, which may include house care, personal care, meal preparation, child care, etc., are distinguished from discretionary activities by the fact that, as suggested earlier, anterior conditions very often stipulate some commensurable time expenditure. Neither are intermediate activities of the same cast as inelastic activities, as evidenced by the multitude of both complementary and substitute goods and services designed to capitalize on that very degree of latitude which individuals may exercise.

This taxonomy of activities corresponding to the dimension of latitude alleviates the indeterminacy of traditional classificatory schemes. Moreover, it is quite efficient as an analytical framework from which to examine the allocation of time.

#### THE ALLOCATION OF TIME

Time, of course, cannot be stored.

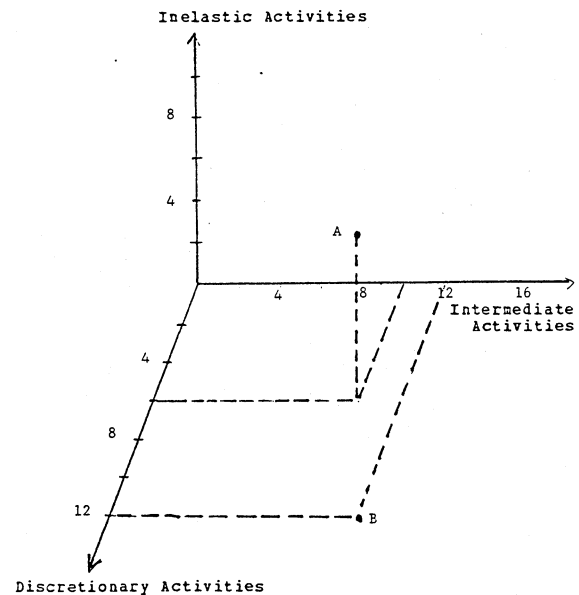
A unique property of time as an indicator is that if time is saved or lost on one activity the time devoted to other activities will either increase or decrease as a result. (Robinson, 1977)

If the exhaustiveness of the proposed taxonomy may be assumed - that is, any activity in which one may engage can be classified into one of the three categories - then the various allocations of a period of time, say, 24 hours, may be portrayed graphically in a space the dimensions of which are the time inputs to inelastic, intermediate, and discretionary activities, as shown in Figure 1. Individual A, for instance, is one whose day is consumed largely by inelastic and intermediate activities, while B, free of inelastic activities, devotes

more time to both discretionary and intermediate activities.

FIGURE 1

THE SPACE IN WHICH TIME IS ALLOCATED



Furthermore, as the time required by inelastic activities increases (decreases), the time allocated to the set of elastic activities - comprised of intermediate and discretionary activities - must decrease (increase) by a commensurate amount. Corresponding to any level of inelastic activities, a time-budget constraint connects the combinations of time expenditures to discretionary and intermediate activities which exhaust the balance of time remaining after the time input to inelastic activities is subtracted from total time. For instance, one with no inelastic activities has 24 hours which must be allocated to some combination of discretionary and intermediate activities. One whose inelastic activities require 8 hours per day has 16 hours which must be allocated to discretionary and intermediate activities.

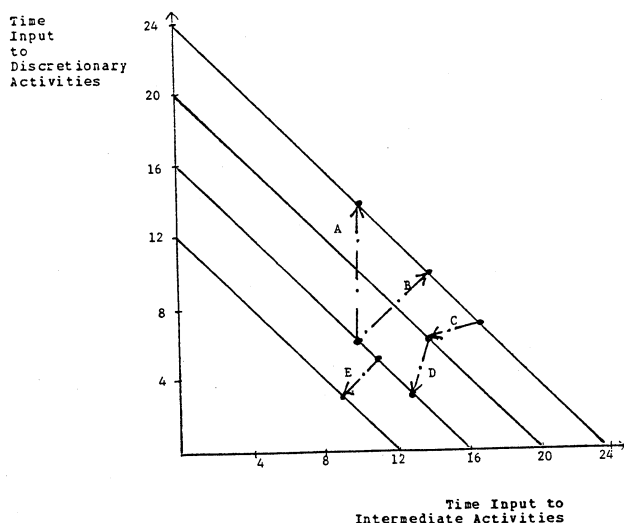
It may be argued that the allocation of time to elastic activities within the time-budget constraint is of most significance, since these are the activities characterized by at least some degree of latitude (in the short run). This is the position taken here. Thus, we treat the time-budget constraint as a given which can only be altered in the short-run by some unusual event, e.g., a change in jobs, retirement, etc., and focus on the allocation of time to discretionary and intermediate activities.

This is not to say that a shift in the time-budget constraint is irrelevant. Indeed, an issue of fundamental importance is the impact of a relaxation (or, conversely, a contraction) of the time-budget constraint on individuals' patterns of time allocation, or stated in more conventional terms, on individuals' choices of activities. Figure 2 shows a number of quite plausible responses. A and B represent somewhat dissimilar responses to a relaxation of the time-budget constraint (from 16 to 24 hours, perhaps brought on by retirement). In B, the individual devotes equal amounts of his increment of 8 hours to both discretionary and intermediate activities (perhaps enabled by a simultaneous move to a retirement community). In C, the individual

<sup>3</sup>As Chapin (1974) proposed, though he subsequently classified activities as either obligatory or discretionary.

responds to a contraction of four hours (perhaps brought on by the advent of a part-time job) by reducing the time devoted to intermediate and discretionary activities by 3 and 1 hours, respectively. The contraction of 4 hours in D, perhaps due to the individual moving from a part-time to a full-time job, has a different impact, however, with discretionary activities reduced more so than intermediate activities. An individual who takes a half-time job, in addition to the full-time job he already has, as depicted in E, may respond similarly.

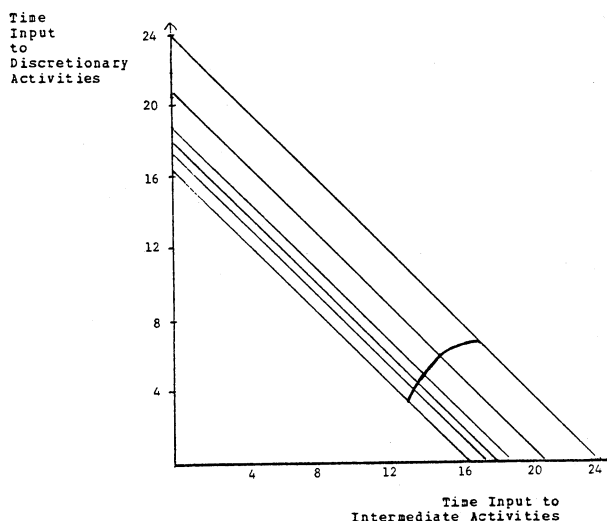
FIGURE 2  
RESPONSES TO RELAXATIONS AND CONTRACTIONS  
IN THE TIME-BUDGET CONSTRAINT



Since the latitude an individual has in allocating some minimum of time to intermediate activities is governed by anterior conditions, as proposed earlier, we might expect intermediate activities to become less elastic relative to discretionary as the time budget constraint becomes increasingly restrictive, as D and E in Figure 2 would suggest. As the time-budget constraint becomes increasingly relaxed, however, individuals exercise increasing latitude in their allocation of time to either intermediate or discretionary activities. With the structural changes occurring throughout society - women entering the work force, four-day weeks, early and postponed retirements, etc. - individual and aggregate responses to changes in time-budget constraints take on increased significance. These responses would best be examined longitudinally. (See Hawes, 1977, for an alternative approach) Cross-sectional data drawn from a University of Michigan Institute for Social Research study (Robinson, 1977, pp. 35, 53), reclassified into the categories developed here, may provide some insight. Women with different levels of inelastic activities showed the patterns of time allocation depicted in Figure 3.

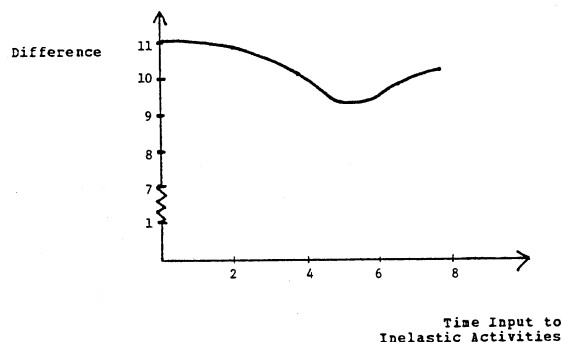
It is instructive to examine these data from a slightly different perspective. Figure 4 depicts the difference between the time allocated to intermediate activities and the time given to discretionary activities as it varies with the time taken up by inelastic activities. The initial decrease is due to a larger contraction in intermediate activities than discretionary activities, which supports the contention that, at least beyond some minimum, there is considerable latitude in the time input to intermediate activities. The upturn in Figure 4 is due to the time input to intermediate activities declining less rapidly than the time input to discretionary

FIGURE 3  
PATTERNS OF TIME ALLOCATION  
BY WOMEN UNDER DIFFERENT TIME-BUDGET CONSTRAINTS



activities, which supports the premise that intermediate activities become less elastic as the time budget constraint become more restrictive.

FIGURE 4  
DIFFERENCE BETWEEN TIME INPUTS TO  
INTERMEDIATE AND DISCRETIONARY ACTIVITIES  
AT DIFFERENT LEVELS OF INELASTIC ACTIVITIES



Despite considerable methodological advances in the study of time (see Juster, forthcoming), reliable data, especially longitudinal data, are difficult and expensive to gather, a factor which will undoubtedly continue to inhibit exploration of the issues posed above. Examination of factors associated with existing patterns of time allocation may be revealing, though.

It should be obvious that consumers with the same amount of time to allocate to either discretionary or intermediate activities may do so quite dissimilarly. As suggested earlier, certain anterior conditions are likely to generate diverse patterns of time (and therefore, money) allocation. Of course, marketers have recognized these associations, as evidenced by the use of such factors as stage in family life cycle, age, etc., as bases for marketing strategies. But, significantly, there is not an absolute correspondence between anterior conditions and time expenditures (just as the correspondence between anterior conditions and money expenditures is not absolute).

For example, while we may expect a positive relationship between size of dwelling and time spent cleaning house, we are likely to observe some individuals residing in smaller homes spending more time cleaning house than others residing in larger homes, even when other anterior conditions which may account for this, e.g., stage in family life cycle, are held constant.

A well specified model of time allocation must incorporate factors which account for these deviations from the time expenditures predicted by anterior conditions. Such mediating factors may include the allocations of time by others, the position of one's time-budget constraint, accessibility, satisfaction, role ideology, and resources, each of which will be discussed in turn.

The allocations of time by others with whom one interacts may mediate one's own time allocation. For instance, one's spouse who does all of the cooking, cleaning, and shopping substantially reduces such time expenditures by that individual. The choices of activities by one's children, peers, etc., likewise govern one's time allocations.

As suggested earlier, the position of one's time-budget constraint is likely to impinge upon one's allocation of time. Highly restrictive time-budget constraints are likely to minimize or eliminate the time input to discretionary activities.

The relationship between accessibility and individuals' allocations of time is well stated by Chapin, from the area of urban planning:

...an activity pattern is contingent not only on a propensity or readiness to engage in that activity, but also on there being an opportunity to engage in that activity in the sense that a facility, service, or other instrumental means is available which permits the activity to take place. (1974, p. 33)

Clearly, the absence of other people with whom to interact severely limits one's time so employed. Similarly, a lack of shopping facilities may limit the time devoted to that activity.

The time devoted to an activity is likely to correspond closely to the relative satisfaction derived from so employing one's time, as Robinson has documented, though the direction of causality is indeterminate (Stone, 1972, p. 181) - e.g., do people allocate their time to those activities which yield the most satisfaction, or do people derive satisfaction from activities on which they spend most of their time. The relationship is strongest, however, between satisfaction and highly discretionary activities, prompting Robinson to conclude:

We suspect, therefore, that even daily routine evolves from a process in which individuals selectively find those activities that are psychologically rewarding and arrange their lives in such a way that participation in these activities can be scheduled more frequently. (1977, p. 191)

Satisfaction is especially likely to distinguish among individuals' time expenditures for activities such as meal preparation, home and auto maintenance, etc., for which goods and services are readily available to substitute for one's time input and which are undoubtedly characterized by a wide range of satisfaction across

people.<sup>4</sup>

Role ideology is a construct which signifies an individual's perception of the degree of correspondence between his or her time input to an activity and the output which results. That is, to what extent does the individual see his or her time input as a necessary prerequisite to some outcome? An oft-cited example in marketing textbooks is that of the early cake mixes which were ill-received supposedly because they required no input on the part of the housewife. Such a role ideology may govern the substitutability of goods and services for the time input to such activities as meal preparation and child care, especially.

The mediating factor of resources provides an individual with the means to exercise a considerable degree of latitude in his or her time allocation. Resources enable one to more readily substitute goods and services (e.g., dishwasher, microwave oven, etc.) for one's time input to certain activities as well as open up a whole range of activities (e.g., photography, boating, gardening, etc.) which call for certain complementary products. "Inferior pursuits" (Linder, 1972, p. 82) - activities for which the time expenditure declines as income rises - are commonly thought to include such activities as housework, though Robinson (1977, p. 69) failed to uncover any such relationship. This does not necessarily mean that the more affluent do not shift their time from less pleasant household activities (say, meal clean-up) to more pleasant ones (say, gourmet cooking). To be more precise, the resource factor most likely interacts with the factor of satisfaction to enable the individual to re-allocate his or her time to relatively preferred activities.

Thus, a more complete specification of the model of time allocation must incorporate both anterior conditions and mediating factors, as depicted in Figure 5.

---

<sup>4</sup> Arndt and Gronmo (1977), examining the time devoted to shopping, suggest that much of the variation which their model failed to capture is likely due to the psycho-social functions performed by stores, e.g., their ability to provide satisfaction to some shoppers more so than others. Also of interest is the development of a social indicator, Lambda, by Hobson and Mann (1975), based on the discrepancy between the amount of time individuals would like to spend in various activities versus the amount of time they actually spend in same.

FIGURE 5

## MODEL OF TIME ALLOCATION

ANTERIOR CONDITIONS		MEDIATING FACTORS	
Age		Allocation of Time by Others	
Health	Probable	Position of Time-	Actual
Stage in FLC	Time	Budget Constraint	Time
Size of Dwelling	Input	Accessibility	Input
Age of Dwelling	to	Satisfaction Derived	to
Size of yard	Activities	from Activity	Activities
Age of Auto(s)		Role Ideology	
		Resources	

## SUMMARY

Activities may be viewed as processes, associated with which are the inputs of time, energy, and goods and services and desired outputs. While the economists have suggested how these inputs are likely to be combined, their model is unsuitable since it ignores constraints posed by anterior conditions on the latitude with which individuals allocate time. A taxonomy of activities corresponding to the degree of latitude one has in determining the time so employed alleviates the indeterminacy of traditional classifications. The allocation of time to the set of elastic activities, e.g., discretionary and intermediate activities, is seen as taking place within some time-budget constraint imposed by the allocation of time to inelastic activities. Finally, mediating factors are likely to account for variation in patterns of time allocation left unexplained by anterior conditions.

## REFERENCES

- J. Arndt, and Sigmund Gronmo. "The Time Dimension of Shopping Behavior: Some Empirical Findings." Proceedings of the Association for Consumer Research Conference (1977): 230-235.
- C. Azzi. "Household Allocation of Time and Church Attendance." The Journal of Political Economy 83-1 (1975): 27-56.
- G. S. Becker. "A Theory of the Allocation of Time." Economic Journal 75 (September 1965): 493-517.
- C. S. Bell. "Value of Time." Social Research 42-3 (1975): 556-563.
- F. S. Chapin, Jr. Human Activity Patterns in the City. New York: Wiley, 1974.
- S. DeGrazia. Of Time, Work, and Leisure. New York: The Twentieth Century Fund, 1962.
- E. Demby. "Psychographics: Who, What, Why, When, Where, and How." in C. W. King and D. J. Tigert, eds. Attitude Research Reaches New Heights. Chicago: American Marketing Association, 1971, pp. 196-199.
- A. Devany. "Revealed Value of Time in Air-Travel." Review of Economics and Statistics 56-1 (1974): 77-92.
- A. W. Evans. "Theory of Valuation and Allocation of Time." Scottish Journal of Political Economy 19-1 (1972): 1-17.
- R. Ferber and L. C. Lee. "The Role of Life Style in Studying Family Behavior." Faculty Working paper 226, College of Commerce and Business Administration, University of Illinois at Urbana-Champaign, 1974.
- S. M. A. Hameed. "Theory of Hours of Work and Nonwork." Indian Journal of Industrial Relations 7-4 (1972).
- D. K. Hawes. "Time Budgets and Consumer Leisure-Time Behavior." Proceedings of the Association for Consumer Research Conference (1977): 221-229.
- R. Hobson and S. H. Mann. "A Social Indicator Based on Time Allocation." Social Indicators Research 1-4 (1974): 439-457.
- J. Jacoby, G. J. Szybillo and C. K. Berning. "Time and Consumer Behavior: An Interdisciplinary Overview." Journal of Consumer Research 2-1 (1976): 320-339.
- F. Thomas Juster, ed. Studies in the Measurement of Time Allocation. Ann Arbor, Michigan: Institute for Social Research, forthcoming.
- S. B. Linder. The Harried Leisure Class. New York: Columbia University Press, 1970.
- B. D. Mabry. "An Analysis of Work and Other Constraints on Choices of Activities." Western Economic Journal 8 (September 1970): 213-225.
- W. Michelson. "Residential Mobility as a Deficit Compensating Process." Paper presented at the Canadian Sociology and Anthropology Association Meetings, Kingston, Ontario, 1973.
- J. N. Morgan. "The Supply of Effort, the Measurement of Well-being, and the Dynamics of Improvement." American Economic Review 48 (May 1968): 31-39.
- F. M. Nicosia and R. N. Mayer. "Toward a Sociology of Consumption." Journal of Consumer Research 3 (September 1976): 65-75.
- F. J. Prochaska and R. A. Schrimper. "Opportunity Cost of Time and Other Socioeconomic Effects on Away-from-home Food Consumption." American Journal of Agricultural Economics 55-4 (1973): 595.
- J. P. Robinson. How Americans Use Time. New York: Praeger Publishers, 1977.

I. V. Sawhill. "Economic Perspectives on the Family." Daedalus 6-2 (1977): 115-125.

P. B. Schary. "Consumption and the Problem of Time." Journal of Marketing 35 (April 1971): 50-55.

P. J. Stone. "Models of Everyday Time Allocation." in A. Szalai, ed. The Use of Time. Paris: Mouton and Co., 1972.

W. D. Wells and S. C. Cosmas. "Life Styles." in R. Ferber, ed. Selected Aspects of Consumer Behavior. Washington, D. C.: NSF RANN, 1977, pp. 299-316.

# INDUCING COMPLIANCE WITH A REQUEST: THE LIST TECHNIQUE<sup>1</sup>

Peter H. Reingen, University of South Carolina

## Abstract

A compliance induction technique where a target is first shown a list of other compliers and then is asked to comply with a request was examined in a field experiment. The technique was found to be effective provided that the number of other compliers was sufficiently large. A formulation based on informational social influence is suggested to account for the findings.

## Introduction

A growing number of investigations have recently focused on variables that affect a person's willingness to comply with a request. Studies have been conducted to examine the effects on compliance of such factors as prior compliance with a small request (e.g., Cann, Sherman, and Elkes 1975; Freedman and Fraser 1966; Reingen in press; Reingen and Kernan 1977; Scott 1977); prior non-compliance with an extreme request (Cialdini, Vincent, Lewis, Catalan, Wheeler, and Darby 1975; Cialdini and Ascani 1976; Reingen in press); cost (e.g., Wagner and Wheeler 1969); need (e.g., Wagner and Wheeler 1969); incentives (e.g., Scott 1976); commitment then cost (Cialdini, Cacioppo, Basset, and Miller in press); legitimization of paltry contributions (Cialdini and Schroeder 1976; Reingen 1977; Reingen in press); and observation of a complier (e.g., Bryan and Test 1967; Wagner and Wheeler 1969).

The last of these factors, observation of someone complying with a request (i.e., a model), seems at first glance the least promising from a pragmatic viewpoint. To enhance the probability of a target's yielding to a request via his observation of a model's behavior would indeed appear to be a rather cumbersome and probably inefficient technique in many large-scale request contexts of interest to consumer behaviorists, such as fundraising drives, blood drives, political campaigns, brand marketing, and the like. Yet, there are techniques employed in everyday compliance settings that bear more than a superficial resemblance to modeling procedures: when a textbook publisher provides professors with a list of new book adopters, when names of contributors to a charity drive are announced during a telethon, and when people are asked to add their name to a petition, to mention but just a few.

The common ground between the everyday practices and the modeling procedures in social psychology appears to be that both are presumably predicated on the efficacy of informational social influence; namely, that a target often uses the behavior of another (or others) to assist him in determining his own action (Deutsch and Gerard 1955). What primarily distinguishes the two is that only the cognate modeling research in social psychology involves a target's observation of a present other's compliance with a request; i.e., the modeling effect appears to be mediated by direct informational social influence. In the case of the common practices, a requester merely provides cues to a target that others have complied; i.e., a compliance effect, if obtained, may be mediated only

indirectly by informational social influence.

But does a technique where a target is merely informed of other compliers by a requester really work, or have practitioners, in the absence of hard evidence, deluded themselves as to the compliance-producing power of the technique? If it does work, is informational social influence the underlying mediator of the effect?

To provide answers to these questions, a study was conducted in a naturalistic setting. Targets (college students) were asked to donate money to the Heart Association in one of ten ways. In the first (the control) a direct request for a donation was made. In the next four conditions subjects were first shown a list of (fictitious) other contributors and then were asked for a donation. Two factors were manipulated: Size-of-donations from others and number-of-donors.

## Informational Social Influence Predictions

It has been well documented that information dependency on others tends to increase with uncertainty (King 1975). Since students are not the typical target of fund-raising efforts, their experienced uncertainty with regard to how much to give should be substantial enough to have them pay especially close attention to the behavior of relevant others (i.e., other's donations). Therefore, the study varied the size of (fictitious) donations of others (low/high). Based on the above, the informational social influence hypothesis would predict that subjects exposed to high donations of others would give a higher amount on the average than subjects informed of others' low donations.

Aside from its conceptual aspect, another purpose of the size-of-donation manipulation was to examine the possibility that the list technique could also be employed to increase the amounts of compliance compared with control outcomes. The key emphasis in previous compliance research centered on the proportion of subjects complying with a request. Yet, it is clear that in many everyday compliance settings, such as charity drives, a second variable of substantial practical interest is the amount of compliance. Thus, based on previous findings (Reingen in press; Reingen 1978), the high donations of other compliers were constructed so that they were greater than what would be normally expected for control subjects.

The study also assessed the effect the sheer number of other compliers has on a target's yielding to a request. Generally speaking, as more people agree on a given behavior, the greater the probability of informational social influence (King 1975). The informational social influence hypothesis would, therefore, suggest that subjects exposed to a long list of donors (12 donors) would be significantly more likely to contribute than subjects exposed to a short list (4 donors), and would be more likely to contribute than control subjects. Although small group research would argue that a target's likelihood of compliance increases with the number of sources of information up to the point of only four sources (Asch 1951), that research is of little relevance here because of its face-to-face interaction basis.<sup>2</sup>

When the variables size-of-donations and number-of-donors are factorially combined, the above suggests that the total compliance effects (i.e., total donations) would be strongest in the high donations/long list cell and weak-

<sup>1</sup>The author would like to thank the Heart Association for its support. This paper reports a portion of a larger study.

## Normative Influence and Hybrid Strategies

A further major purpose of the field experiment was to incorporate into the experimental design a procedure based on normative social influence so that the differential effects of both types of social influence (informational and normative) could be assessed as compliance induction techniques.<sup>3</sup> Informational social influence is largely self-inflicted and occurs when another's behavior is attended to and has impact as the result of the receiver's attempting to be "correct" or "appropriate," e.g., when a person's giving to charity is primarily a function of his perception of the giving behavior of others. Generally, informational social influence is unintended by the source, as is presumably the case in the present study where the fictitious donors are hypothesized to be the primary source of influence.

In contrast, normative social influence is largely instrumental and results when a receiver accepts influence in order to gain some reward beyond merely being appropriate. Normative social influence is generally intended by the source who is perceived as a mediator of rewards. Frequently, however, people can influence others to do things in normative social influence situations without mediating specific rewards (or punishments). That is, people often comply with a request because they find it intrinsically self-rewarding. To achieve this form of normative social influence, a requester provides cues, or stimuli, that are designed to make salient certain pre-established dispositions acquired during socialization, such as the internalized norm of social responsibility, which direct behaviors the target finds self-rewarding. In the context of the present study, this could mean that common additions of such cues as "We need your help" or "You should give to a good cause" may be sufficient enough to elicit greater compliance. However, such simple moral exhortations have not been found to be effective (Bryan and London 1970; Darley and Betson 1973). Cues are apparently needed that lead a target into believing that his own behavior will have an important bearing on another's welfare; i.e., cues that activate the norm of personal responsibility (e.g., Schwartz 1970). This was attempted in the present study by incorporating to the experimental script a sentence that in effect told targets that their help was needed to prevent heart attacks for people they might even know. It was predicted that subjects exposed to this cue would be more likely to contribute than control subjects.

In the final four conditions, the study assessed the compliance effects of four hybrid strategies. Subjects were first subjected to the normative social influence attempt, then were shown a list of fictitious donors, with the number of donors and size of donations varied as previously described, and finally were asked for a donation. Since the hybrid strategies would appear to reap the benefits of both normative and informational social influence attempts, they were expected to produce greater compliance with the donation request than their component compliance techniques.

<sup>2</sup>Evidence provided by Gerard, Wilhelmy, and Connolly (1968) suggests that when interaction is less direct, conformity tends to be enhanced as the size of the majority increases from two to seven. Since the present study involved indirect observation contexts, a number of compliers larger than seven was judged to be sufficient for the effect to materialize.

<sup>3</sup>The following discussion is largely based on King (1975).

## Subjects

Subjects were 300 male students at the University of South Carolina. Only those students qualified who were walking or sitting alone along university walkways during the hours of 10:00 a.m. to 5:00 p.m., and no subjects known to an experimenter were selected.

## Procedure

The experimenters, three male college students, had been thoroughly instructed and they were equipped with the identification badges, information brochures and donation envelopes commonly employed in fund-raising efforts by the Heart Association. An interaction was initiated by an experimenter's introduction of himself as representing the Heart Association. After the common introductory remarks, subjects were randomly assigned to one of ten conditions, 30 subjects each, according to a pre-specified treatment schedule that varied across experimenters. The schedule was constructed so that an experimenter completed exactly 10 replications per condition. After an interaction had been completed, an experimenter contacted the next subject that qualified.

In the first condition, the request-only control, the experimenter stated, "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?" In the next four conditions, the list-then-request conditions, a subject was first shown a list of fictitious donors with their "donations" and then was asked for a donation. Specifically, the experimenter said, "As part of our annual campus fund-raising drive, I'm collecting money for the Heart Association. (While the experimenter stated the following, he showed the subject a list of donors of both sexes with their donations.) As you can see, other students have given a donation already. (The experimenter then counted to two and continued.) Would you be willing to help also by giving a single donation?" The length of the list (short--four donors/long--twelve donors) and the size of donations (low/high) were varied. The low donations had a mean of \$.25 (range \$.15 to \$.35) for both short and long lists, whereas the high donations averaged to \$.85 (range \$.75 to \$.95). The names of donors were, of course, the same across the donation size factor. Thus, three identical sets, one for each experimenter, with four lists each were utilized.<sup>4</sup>

In the sixth condition, the normative influence-then-request condition, the experimenter stated, "As part of our annual campus fund-raising drive, I'm collecting money for the Heart Association: We need your financial support so that you can help us in preventing heart attacks for people you might even know. Would you be willing to help by giving a single donation?"

The final four conditions, the hybrid strategy conditions, were the same as conditions two through five, except that subjects were shown the lists of donors after they had been told that their financial support was needed.

<sup>4</sup>The lists were constructed by asking acquaintances of the author to write down their name and a predetermined donation amount, without collecting any funds. It is extremely improbable that subjects knew any of the "donors." Subjects were of course not permitted to add their name to the list. A blank page was attached to the list for subjects that wanted their name and donation listed. Very few of the subjects expressed that desire throughout the study.



## Results

Table 1 presents the results for frequency of donation, average amount donated, and total amount donated.

TABLE 1  
DONATION RESULTS

Condition <sup>a</sup>	Percent of Compliance	Donation	
		Average	Total
1. Request-only control	40%	\$.28	\$ 3.35
2. Short/low list-then-request	47	.41	5.75
3. Long/low list-then-request	63	.37	6.99
4. Short/high list-then-request	47	.46	6.48
5. Long/high list-then-request	60	.55	10.04
6. Normative influence-then-request	53	.44	6.99
7. Normative influence, short/low list-then-request	57	.35	6.07
8. Normative influence, long/low list-then-request	67	.33	6.63
9. Normative influence, short/high list-then-request	63	.57	10.84
10. Normative influence, long/high list-then-request	70	.68	14.22

<sup>a</sup> Base is 30 in each condition.

### Compliance Frequencies

An initial Chi Square analysis on frequency of donation within each condition showed no significant differences between experimenters, with levels of significance ranging from .38 to .87. Hence, the compliance data were quite free of experimenter effects, and the subsequent analyses could therefore be performed on collapsed data.

The hypotheses with regard to compliance frequencies seemed most appropriately tested with a series of four planned orthogonal contrasts.

The first compared the combination of the long list conditions (conditions 3,5,8, and 10) with the combination of the short list conditions (conditions 2,4,7, and 9) and the control (condition 1). The informational social influence prediction that subjects exposed to a long list of compliers would be significantly more likely to contribute than subjects exposed to a short list or no list was clearly confirmed ( $X^2 = 5.02$ ,  $d.f.=1$ ,  $p < .02$ ).<sup>5</sup>

The second orthogonal comparison between the control and the combination of the four short list conditions was insignificant ( $X^2 = 1.22$ ,  $d.f.=1$ ,  $p > .10$ , two-tailed), suggesting that the compliance effects are obtained only if the number of other compliers is sufficiently large.

<sup>5</sup>All 2 X 2 contingency analyses were corrected for continuity and all directional tests were one-tailed (Siegel 1956). Since the comparisons were orthogonal, the error rates were not inflated.

Thus, when coupled with the results of the first comparison, the findings suggest a length-of-list effect.

In contrast, no size-of-donations effect on compliance rates was observed, as is attested to by the third orthogonal contrast between the combination of the low donation conditions (conditions 2,3,7 and 8) and the combination of the high donation conditions (conditions 4,5,9, and 10) which was insignificant ( $X^2 = .02$ ,  $d.f. = 1$ ,  $p > .10$ , two-tailed).

It was expected that the hybrid strategies (conditions 7 through 10) would induce greater compliance than the list-only conditions (conditions 2 through 5). The fourth orthogonal comparison involving these conditions provided only marginal support for the prediction ( $X^2 = 2.09$ ,  $d.f. = 1$ ,  $p < .08$ ).

The hypothesis that the normative influence condition (condition 5) would produce greater compliance than the control (condition 1) could not be directly assessed by the series of orthogonal contrasts employed. Thus, this comparison was performed separately, with the unexpected result that no significant difference was found ( $X^2 = .60$ ,  $d.f. = 1$ ,  $p > .10$ ). Conditions 2 through 5 and conditions 7 through 10 were also analyzed separately in an attempt to determine interaction effects between length-of-list and size-of-donations on compliance. The resulting Chi Squares were insignificant ( $p$ 's  $> .10$ ), thus obliterating the need for a partitioning of variance via an arc sin transformation of the proportions (Rao 1952).

### Compliance Amounts

Due to the presence of unequal cell frequencies of com-

pliers (i.e., nonorthogonality) and two no-list conditions, the hypotheses with regard to donation amounts could not be tested by standard approaches for the factorial analysis of variance. Instead, a linear model appropriate for a 2 (donations: low/high) X 2 (length: short/long) X 2 (normative influence: absent/present) design that included a variable for the two no-list conditions (coded 1 if the response came from the control and -1 otherwise) was constructed (Mendenhall 1968). The importance of each effect was then assessed via general linear hypothesis procedures (e.g., Perreault and Darden 1975). Using regression analysis, simple models not including a particular effect were compared with complete models which included the effect. The results showed that only the size-of-donations factor had a significant effect on the data. Thus, the other factors could be ignored, and the analysis of the remaining one-factor model suggested a highly significant effect of size-of-donation of other compliers on the dependent measure ( $F = 17.61$ ,  $d.f. = 1/168$ ,  $p < .01$ ). The overall means ( $\bar{x}_{low} = .37$  versus  $\bar{x}_{high} = .57$ ) indicate that the results were in the predicted direction.<sup>6</sup>

#### Compliance Totals

Given this pattern of findings, it is not surprising that along the practical dimension of total funds obtained, the results are consistent with the informational social influence hypothesis. The conditions where long/high lists were involved (conditions 4 and 10) produced greater donation totals (\$10.04 and \$14.22, respectively) than their short/low list counterparts (\$5.75 for condition 2 and \$6.07 for condition 7). When compared with the control outcome, from a practical viewpoint the results are even more impressive. The long/high list-then-request condition produced a total three times as high as that of the control, and the normative influence, long/high list-then-request condition produced 4.2 times that of the control.

#### Discussion

These findings are consistent with the informational social influence hypothesis, but they do not ultimately confirm it. Overall, the subjects were consistently more likely to comply with the donation request the greater the number of other compliers, and the subjects offered a consistently higher amount of compliance on the average the higher its level for other compliers. Nevertheless, alternative explanations may exist, especially in view of the fact that manipulation checks were not judged to be feasible in this field experiment.

The data also suggest that to achieve greater compliance a minimum number of other compliers is necessary if informational social influence is to be activated in indirect observation contexts. When the number of other compliers was small (four), the compliance increases tended to be insignificant. The suggestion that in indirect as opposed to direct observation settings more models are needed to pose a sufficient enough challenge to an observer's typical self-image of benevolence and

<sup>6</sup>When an analysis of variance for a completely randomized design was performed on the data, the same overall conclusion emerged. Multiple orthogonal comparisons on the data showed that the high contribution conditions produced greater donations than the combination of the low donation conditions and the control; and that no significant difference existed between (a) the control and the low donation conditions, (b) the short list and long list conditions, and (c) the list-only conditions and the hybrid strategy conditions.

charity appears plausible. The more direct a contact, the greater tends to be the influence potential (King 1975).

No support was found for a compliance technique based on normative social influence. Of course, only one of many possible operationalizations of normative influence was utilized, and its cues may simply have been too weak to produce the desired effects. It is entirely possible that with a different operationalization, one which would have made the norm of personal responsibility more salient, the results would have been more impressive. When used in conjunction with informational social influence, however, the overall compliance proportions were marginally more favorable, suggesting an augmented effect.

There is, of course, a practical utility to the findings as well. The findings suggest an effective approach to increasing the frequency and amount of compliance. Thus, the donation totals for the long/high list conditions were much greater compared with the control outcome, a result that fund-raisers in particular should consider of value. The technique can also be easily implemented, and it can be employed in an ethically responsible fashion with the use of real donors and donations. However, the effects have only been demonstrated in the context of a prosocial request. Although preliminary evidence (Reingen 1978) suggests that the technique may possess cross-situational reliability provided that the norms that govern compliance are not too strongly violated, further research is needed to determine the generality of the effect.

#### References

- S.E. Asch, (1951), "Effects of Group Pressure on the Modification and Distortion of Judgements," in Groups, Leadership, and Men, ed. H. Guetzkow, Pittsburgh, Pa.: Carnegie Press, 177-90.
- J.H. Bryan and M.A. Test, (1967), "Models and Helping: Naturalistic Studies in Aiding Behavior," Journal of Personality and Social Psychology, 6, 400-7.
- J.H. Bryan and P. London, (1970), "Altruistic Behavior by Children," Psychological Bulletin, 73, 200-11.
- A. Cann, S.J. Sherman, and R. Elkes, (1975), "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, 774-82.
- R.B. Cialdini and K. Ascani, (1976), "Test of a Concession Procedure for Inducing Verbal, Behavioral, and further Compliance with a Request to Donate Blood," Journal of Applied Psychology, 61, 295-300.
- R.B. Cialdini, J.T. Cacioppo, R. Basset, and J.A. Miller, (in press) "The Low-Ball Procedure for Producing Compliance: Commitment then Cost," Journal of Personality and Social Psychology.
- R.B. Cialdini and D.A. Schroeder, (1976), "Increasing Compliance by Legitimizing Paltry Contributions: When even a Penny Helps," Journal of Personality and Social Psychology, 34, 599-604.
- R.B. Cialdini, J.E. Vincent, S.K. Lewis, J. Catalan, D. Wheeler, and B.L. Darby, (1975), "Reciprocal Concessions Procedure for Inducing Compliance: The Door-in-the-Face Technique," Journal of Personality and Social Psychology, 31, 206-15.
- J.M. Darley and D.C. Batson, (1973), "From Jerusalem to Jericho: A Study of Situational and Dispositional Vari-

- ables in Helping Behavior," Journal of Personality and Social Psychology, 27, 100-8.
- M. Deutsch and H. Gerard, (1955), "A Study of Normative and Informational Social Influences upon Individual Judgement," Journal of Abnormal Social Psychology, 51, 629-36.
- J.L. Freedman and S. Fraser, (1966), "Compliance without Pressure: The Foot-in-the-Door Technique," Journal of Personality and Social Psychology, 4, 195-202.
- H.B. Gerard, R.A. Wilhelmy, and E.S. Connolley, (1968), "Conformity and Group Size," Journal of Personality and Social Psychology, 8, 79-82.
- S.W. King, (1975), Communication and Social Influence, Reading, Mass.: Addison-Wesley Publishing Company.
- W. Mendenhall, (1968), The Design and Analysis of Experiments, Belmont, Ca.: Wadsworth Publishing Company.
- C.W. Park and V.P. Lesig, (1977), "Students and Housewives: Differences in Susceptibility to Reference Group Influence," Journal of Consumer Research, 4, 102-10.
- W.D. Perreault and W.R. Darden, (1975), "Unequal Cell Sizes in Marketing Experiments: Use of General Linear Hypothesis," Journal of Marketing Research, 12, 333-43.
- C.R. Rao, (1952), Advanced Statistical Methods in Biometric Research, New York, N.Y.: Wiley.
- P.H. Reingen, (1977), "On the Marketing of Fund-Raising Campaigns: Foot-in-the-Door and Legitimization of Paltry Contributions," Working Paper, Center for Marketing Studies, University of South Carolina.
- P.H. Reingen, (1978), "Social Influence as a Compliance Induction Technique," Working Paper, Center for Marketing Studies, University of South Carolina.
- P.H. Reingen, (in press), "On Inducing Compliance with Requests," Journal of Consumer Research.
- P.H. Reingen and J.B. Kernan, (1977), "Compliance with an Interview Request: A Foot-in-the-Door, Self-Perception Interpretation," Journal of Marketing Research, 14, 365-69.
- S.H. Schwartz, (1970), "Moral Decision Making and Behavior," in Altruism and Helping Behavior, eds. J. Macaulay and L. Berkowitz, New York: Academic Press, 127-41.
- C.A. Scott, (1976), "The Effects of Trial and Incentives on Repeat Purchase Behavior," Journal of Marketing Research, 13, 263-69.
- C.A. Scott, (1977), "A Self-Perception Approach to Modifying Socially-Conscious Behavior: The Foot-in-the-Door Technique," Journal of Consumer Research, 4.
- S. Siegel, (1956), Non-Parametric Statistics for the Behavioral Sciences, New York, N.Y.: McGraw-Hill.
- C. Wagner and L. Wheeler, (1969), "Model, Need, and Cost Effects in Helping Behavior," Journal of Personality and Social Psychology, 12, 111-16.

EXTENDING THE SELF-PERCEPTION EXPLANATION:  
THE EFFECT OF CUE SALIENCE ON BEHAVIOR

Carol A. Scott, University of California, Los Angeles  
Alice M. Tybout, Northwestern University

Abstract

A 2 x 2 factorial experiment was conducted in which consumers received either an incentive or no incentive for choosing one brand (focal brand) over another brand (nonfocal brand). Then, they were given several pieces of information about the brands. The order of these pieces of information was varied so that either participants' behavior was the only highly salient cue or multiple cues were salient. Although no significant effect of the treatments was observed on subsequent choice of the focal brand, a highly significant incentive by cue salience interaction was obtained for choice of the nonfocal brand. When behavior was the only highly salient cue, the presence of an incentive operated as a discounting cue resulting in less choice of the nonfocal brand than no incentive. In contrast, when multiple cues were salient, the incentive served as a reinforcing cue, resulting in more choice of the nonfocal brand than no incentive.

Introduction

During the past decade, interest has emerged in understanding how consumers' experiences influence their subsequent attitudes and actions. Much of the inquiry pertaining to this issue has been motivated by self-perception theory (Bem, 1972). According to this formulation, individuals examine their own past behavior and its circumstances in an effort to determine the cause of their behavior. When no plausible external cause for the behavior is present, the behavior is interpreted as a veridical representation of the actors' internal states and a positive attitude toward that behavior is inferred. This inferred attitude, in turn, may guide subsequent behavior. If, however, plausible external causes for the behavior are present (i.e., an incentive is offered for the behavior or there is a lack of volition in performance of the behavior), the role of internal states in determining the behavior is discounted and behavior is likely to have little systematic effect on attitudes and future behavior. A substantial number of social psychological investigations have examined the self-perception premise that individuals use their own behavior and its circumstances to infer their attitudes and/or direct their future actions (e.g., Calder and Staw, 1975; Freedman and Fraser, 1966; Lepper, Nisbett and Greene, 1973; Snyder and Cunningham, 1975; Valins, 1974). Furthermore, consumer researchers have tested the effectiveness of self-perception based strategies in marketing contexts (e.g., Dholakia and Sternthal, 1977; Reingen and Kernan, 1977; Scott, 1976, 1977; Swinyard and Ray, 1977). These investigations have provided support for self-perception predictions. In contrast, several studies have found that self-perception theory predictions are obtained only under certain conditions (e.g., Scott and Yalch, 1978; Tybout, 1978).

This observation has motivated the search for factors that are necessary for self-perceptions to occur. One factor identified is the salience of an individual's own behavior. In a field experiment, Tybout (1978) gained compliance with a behavior under two conditions; for some research participants a discounting cue was present, while for others the discounting cue was absent. The salience of the behavior was manipulated by either asking individuals to identify the causes of their behavior (i.e., high salience), or not asking them to

complete this task (low salience). When behavior was made highly salient, the self-perception prediction was obtained; the presence of the discounting cue undermined internal attribution and persistence of the behavior. In contrast, when behavior was not made highly salient, no self-perception effect was observed.

Tybout's findings suggest that an individual's own behavior must be highly salient if the self-perception process is to occur. Further evidence for this contention emerges in a study reported by Scott and Yalch (1978). They offered some consumers an incentive for choosing to taste a new soft drink while others were not given an incentive for performing this task. Prior to tasting the soft drink, some participants were given an opportunity to visually examine the drink, while others were not. They found that when individuals were allowed to usually examine the beverage before tasting it, the incentive served as a discounting cue. Product evaluation was more favorable when the incentive was absent than when it was present. In contrast, when no opportunity was given to visually inspect the product selected, the incentive resulted in a more favorable evaluation than when there was no incentive for product choice.

The Scott-Yalch findings can be explained by interpreting the variation in opportunity to examine the product visually as a salience induction. When research participants were given the opportunity to visually inspect the product chosen, their past behavior was highly salient. As a result, they engaged in an attributional analysis of that behavior. When no visual examination opportunity was provided, subjects' own behavior was not salient. Thus, no attributional analysis was undertaken. Instead, cues about the product were simply processed and the positive incentive cue, like other positive cues, served to enhance or reinforce favorable evaluations of the product. Apparently, the effect of both behavior and other cues in the environment, such as incentives, may be mediated by the salience of behavior.

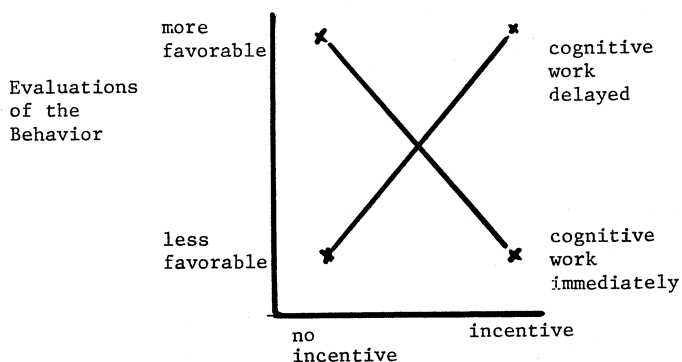
In sum, the evidence reviewed underscores the fact that the salience of individuals' own behavior is a necessary condition for attributional analysis to occur. However, it does not provide an explanation of the mechanism by which salience mediates attributions. This issue is addressed by Pryor and Kriss (1977). They manipulated the timing or sequencing of different attribution cues and observed that the attributions subjects made were related to the relative availability of the cues from memory. Highly available cues guided the attributions made. The implication is that cue salience mediates information processing by influencing the extent to which various cues are available to an individual for acquiring a disposition.

The purpose of the present study is to investigate further the process by which cue salience mediates influence. This entailed having all subjects engage in a behavior. Some subjects received an incentive for performing the behavior, while others did not. Next, all subjects received two pieces of information relevant to their behavior. One was opposed to the behavior. It was intended to motivate subjects to

process the information they had pertaining to the object of behavior. The other piece of information provided support for the behavior subjects had engaged in. The timing of these two pieces of information was manipulated so that subjects assigned to the cognitive work immediate treatment were presented with information that opposed their behavior, immediately after they had engaged in the behavior. This induction was intended to make subjects' own behavior salient. In the cognitive work delayed condition, information opposing subjects' behavior was deferred until after they had received information that was consistent with their behavior. In this situation, cues beyond subjects' own behavior should become salient.

The design of this study is a 2 x 2 factorial, with 2 levels of incentive (no incentive, incentive) and 2 levels of timing (cognitive work immediately, cognitive work delayed). The predicted results are shown in Figure 1. When cognitive work was stimulated immediately after the behavior, it was hypothesized that the salience of that behavior would cause the self-perception process to occur. In the no-incentive condition, response would be more favorable than in the incentive condition, because the incentive would serve as a discounting cue for internal attribution of the behavior. In contrast, when cognitive work was delayed until after additional information was obtained, no causal analysis of behavior was expected because multiple cues would be salient. A more favorable response was expected in the incentive condition than the no-incentive condition because the incentive would serve to reinforce behavior (Scott and Yalch, 1978). Further, it was predicted that within the no-incentive condition, stimulating cognitive work immediately after behavior would lead to more favorable evaluations and behavior than delaying cognitive work until after further information had been received. This was expected because, in the no-incentive condition, behavior should be a positive cue and thus circumstances which increase the salience of this cue (i.e., cognitive work immediately) should lead to more favorable evaluations. Finally, in the incentive condition, delaying cognitive work until further information had been received was expected to produce more favorable evaluations and behavior than stimulating cognitive work immediately after behavior. This result was anticipated because the delay would reduce the salience of behavior leading the incentive to be interpreted as a reinforcing rather than discounting cue.

FIGURE 1  
PREDICTED INTERACTION



Participants were 174 female residents of Columbus, Ohio, between the ages of 21 and 60, who were recruited by a commercial firm for a marketing research study. Most of them were members of volunteer or service organizations which received a \$2 contribution from the research firm for each person participating. The women were scheduled to participate in the study in groups of 5-10

and these groups were randomly assigned to one of the four experimental treatments. Tables with side panels were used to insure that women in the same group could not see each others' responses or communicate with each other during the study. Twenty women were eliminated from the final analysis either because they failed to complete the questionnaire or because they chose the nonfocal brand.<sup>1</sup>

#### Procedure

Upon arrival at the research firm, groups of participants were escorted to a test kitchen and were introduced to two experimenters. One experimenter explained that the purpose of the study was to obtain consumers' evaluations of two new formulations of a diet soft drink for a well known manufacturer. Participants were told that they would receive several pieces of information about these new brands, including a sample of one of the brands to taste, as a basis for making their evaluations. Then, they were instructed to turn over the booklet which had been placed on the table in front of them and read descriptions of each of the brands which purportedly had been taken from advertising campaigns being developed for them. At this point the procedure varied depending upon the incentive treatment to which the participants had been assigned.

No Incentive Condition. Individuals assigned to this treatment were simply instructed to read the two brand descriptions carefully and then select the brand they wished to sample. For purposes of control, it was important that all individuals voluntarily select the same brand for trial. Therefore, the description of one brand, which we shall refer to as the focal brand, was somewhat more favorable than that for the other nonfocal brand. Nearly all (90%) participants in this treatment chose the focal brand.

Incentive Condition. Individuals assigned to this treatment also were told to read the two brand descriptions carefully and then select the brand they wished to sample. In addition, they were informed that in order to encourage trial the focal brand, on that particular day, a coupon worth 50¢ at a popular fast-food restaurant would be given to those who decided to sample that brand. Ninety-six percent of the participants in this condition chose the focal brand.<sup>2</sup> Immediately after participants in this treatment made their choice, one of the experimenters distributed coupons to those who had chosen the focal brand, thereby eliminating any delay in receiving the reward.

Participants were asked to indicate their choice by marking it on their questionnaire. The rationale for this task given to research participants was that this information would facilitate data processing. In reality, it served to reinforce the fact that a decision had been made.

After participants selected the brand they wanted to sample, they turned the page in their booklets and answered some general questions about the product category. Next, individuals received further information about the two brands which included a manipulation of the second independent variable, timing of cognitive work.

1. There was no significant difference in subject attrition for the various treatments.

2. There was no significant difference in the number of participants choosing the focal brand in the no-incentive and incentive treatments. All individuals who did not choose the focal brand were deleted from the analysis.

Cognitive Work Immediately. At this point, individuals assigned to the cognitive work immediately treatment received information designed to motivate cognitive processing regarding their evaluation of the focal brand. Specifically, processing was stimulated by providing them with information which was inconsistent with their choice behavior (i.e., information which indicated that the nonfocal brand was more desirable than the chosen, focal brand). Participants in this treatment heard the experimenter say:

"Before you taste a sample of the brand you have chosen, we thought you might find it interesting to learn how others like yourself have responded to these brands in previous research. Several months ago we ran some tests in which individuals were given both brands to taste and compare. In these studies we found that more than 80% of the people preferred brand 534 (the nonfocal brand) to brand 453 (the focal brand)."<sup>3</sup>

This manipulation was felt to be an appropriate way to motivate cognitive processing since it created uncertainty and uncertainty has been found to motivate information search (Sears and Freedman, 1967). After receiving this negative information about the focal brand, individuals were allowed to gather further data by tasting the focal brand. Pretesting of the drink used as the focal brand indicated that most individuals evaluated the taste of the drink quite favorably. Therefore, this second piece of information was considered to be positive or supportive of the initial choice behavior.

Cognitive Work Delayed. Individuals assigned to the cognitive work delayed treatment received the same positive and negative information about the focal brand as those in the cognitive work immediately treatment. However, the ordering of these two pieces of information was reversed to delay cognitive processing until after further information about the focal brand was obtained (i.e., the negative information followed, rather than preceded, the positive taste information).

Following this information, participants evaluated the focal brand by responding to attitudinal questions in their booklets. Then, individuals were informed that a lottery, in which the prize was a case (24 cans) of diet soft drink, would be held for all study participants. They were asked to mark on their booklet the number of units of each of the new brands (focal and nonfocal) and/or a popular national brand they would want in their case of diet soft-drink if they won the lottery.

Finally, the experimenters led an informal discussion of the study with participants in which they answered questions and attempted to uncover any suspicion about the study purpose and procedure. No evidence of suspicion was found. In general, participants appeared to accept the cover story. The majority of their questions focused on issues such as the identity of the manufacturer and when the brands might be introduced in the marketplace. At the end of the discussion, participants were asked to refrain from discussing the study until the research was complete (two weeks) and were thanked for their cooperation.

#### Dependent Measures

Both attitudinal and behavioral dependent measures were taken. However, due to space limitations and the fact that attitudinal data analysis was not complete in time to be included in this paper, only the measure used to check random assignment and the behavior measure are presented and discussed here.

3. Three digit numbers were used to identify the two brands in the study to avoid any confounding effect of a brand name or letter.

Affect Toward Product Category. At the outset of the study participants expressed their affect toward the product category by responding to the question: Assume that you were to taste a sugar-free soft drink. How likely is it that you would enjoy it? Individuals drew a vertical line through the 10-inch line scale, ranging from not at all likely to extremely likely, at the point which best expressed their opinion. The responses were later scored using a ruler. This measure was used to check the random assignment of groups of participants to treatments. If the random assignment procedure was successful, no difference between treatments on this measure should be obtained.

Choice of Focal Brand. The number of cans of the focal brand that each individual indicated that she would want in her case of diet soft drink, if she won the lottery, served as a measure of choice of the focal brand.

Choice of the Nonfocal Brand. The number of cans of the nonfocal brand that each participant indicated that she would want in her case of diet soft drink, if she won the lottery, served as a measure of choice of the nonfocal brand.

#### Results

##### Randomization Check

Analysis of participants' affect toward the product category by treatment revealed no systematic differences between treatments on this measure (F's for all factors and interactions < 1, n.s.). Therefore, the random assignment procedure is judged to have been successful.

##### Behavior Measures

The impact of the treatments on choice of the focal and nonfocal brand was assessed using the BMDP program for analysis of variance. Affect toward the product category was employed as a covariate in all analyses.<sup>5</sup> It was anticipated that the interaction of incentive and timing of cognitive work shown in Figure 1 would be obtained for the focal brand.

Counter to our hypotheses, an analysis of variance of the focal brand data in Table 1 resulted in no significant effects. However, an analysis of variance on choice of the nonfocal brand (Table 1) did yield a highly significant incentive by timing of cognitive work interaction [ $F(1,169) = 17.52, p < .001$ ].<sup>6</sup>

Furthermore, systematic contrasts between treatments for choice of the nonfocal brand reveal that the pattern of data is the one which had been anticipated for the focal brand (Table 2). Thus, the predicted effects of the treatments on choice were not obtained for the focal brand, but did appear for the nonfocal brand.

4. The number of cans of the national brand chosen was not analyzed as a dependent measure in this study because doing so would use up all the degrees of freedom on the choice measure.

5. Affect toward the product category was an appropriate covariate for this analysis because it was not influenced by the treatments (see discussion under randomization check), but it did contribute significantly to the high variance in each treatment condition. Use of the variable as a covariate did not change the pattern of the results, but did increase the significance of the results in some cases.

6. The covariate was not significant in this particular analysis ( $F < 1$ ).

TABLE 1  
CHOICE OF FOCAL AND NONFOCAL BRANDS  
BY TREATMENT

		$\bar{X}$	s			$\bar{X}$	s
No Incentive	Focal Brand	12.42	9.26	Focal Brand	11.83	9.29	
	Non-focal Brand	6.87	6.02	Non-focal Brand	2.62	6.04	
		n=42				n=46	
		$\bar{X}$	s			$\bar{X}$	s
Incentive	Focal Brand	15.84	9.28	Focal Brand	12.83	9.26	
	Non-focal Brand	2.94	5.96	Non-focal Brand	6.31	6.02	
		n=44				n=42	
		Immediate		Delayed			
		Timing of Cognitive Work					

TABLE 2  
PLANNED CONTRASTS FOR SIGNIFICANT INCENTIVE BY  
TIMING OF MOTIVATION FOR COGNITIVE WORK INTERACTION

Contrast <sup>a</sup>	Significance <sup>b</sup>
1. NI-CI and I-CI	t=3.04 p<.001
2. NI-CD and I-CD	t=-2.88 p<.005
3. NI-CI and NI-CD	t=3.32 p<.001
4. I-CI and I-CD	t=-2.60 p<.005

<sup>a</sup>NI-CI = no incentive, cognitive work immediately  
NI-CD = no incentive, cognitive work delayed  
I-CI = incentive, cognitive work immediately  
I-CD = incentive, cognitive work delayed

<sup>b</sup>One-tailed tests were performed because the direction of the treatment differences were predicted. Further, since the incentive by timing of cognitive work interaction was significant for the choice of the nonfocal brand, contrasts involved protected t's and no correction for error rate was needed (Winer, 1971). Degrees of freedom for all contrasts = 169.

#### Discussion

The major finding of the study is the significant incentive by timing of cognitive work interaction which was obtained for choice of the nonfocal brand. Specifically, it was found that:

1. When cognitive work was stimulated immediately after the choice behavior, more choice of the nonfocal brand resulted under the no-incentive condition than under the incentive condition.
2. When cognitive work was delayed until after further

information was obtained, more choice of the nonfocal brand was produced under the incentive condition than under the no-incentive condition.

3. When no incentive was offered for initial choice of the focal brand, more choice of the nonfocal brand resulted under the cognitive work immediately condition than under the cognitive work delayed condition.

4. When an incentive was offered for the initial choice of the focal brand, more choice of the nonfocal brand was produced under the cognitive work delayed condition than under the cognitive work immediately condition.

In contrast, the predicted interaction for the focal brand was not observed.

One plausible explanation for the results observed exists. Prior to the lottery, individuals' evaluations of the focal brand were independent of the nonfocal brand (i.e., they gave ratings of their liking for the focal brand, the degree to which the focal brand possessed various attributes, etc. without reference to the nonfocal brand). Under these conditions, it seems clear that the predicted incentive by timing of cognitive work interaction should be obtained for evaluations of the focal brand. However, the lottery measure of choice forced a dependent relationship between behavior with respect to the focal and nonfocal brands.

Furthermore, when this measure was taken, individuals had two major pieces of information they could use to make their choice: (1) their attitude toward the focal brand, which appears to demonstrate the predicted incentive by timing of cognitive work interaction (see footnote 7), and (2) knowledge that others like themselves preferred the nonfocal brand to the focal brand. In view of the dependency between choice of the focal and nonfocal brands and the nature of the information available, it is not surprising that the nonfocal brand exhibits the predicted interaction. Individuals who were relatively positive about the focal brand reasoned that they would like the nonfocal brand even more and therefore chose it, while those who evaluated the focal brand less favorably had little reason to believe that they would like the nonfocal brand much better and therefore were less inclined to choose it. Thus the effect of the treatments on attitudes toward the focal brand is observed on choice of the nonfocal brand.

In summary, this study makes several contributions. It replicates previous research (Scott and Yalch, 1978), demonstrating that effects of behavior and behavior-contingent incentives on subsequent evaluations of a product are mediated by the salience of behavior. When behavior is highly salient, an incentive serves as a discounting cue for internal attribution of the behavior, resulting in less favorable evaluations than when no incentive is given, while when behavior is not highly salient, an incentive serves as a reinforcing cue enhancing evaluations relative to no incentive. Further, the present study implicates cue availability as the process by which salience determines the effect of behavior in subsequent cognitive processing. Finally, the results of this research suggest that when brand choices are interdependent and direct comparisons are made (i.e., side by side test in advertisements), self-perceptions of behavior with respect to one brand may influence subsequent choice of a competitive brand.

7. In fact, preliminary results of analysis of attitudinal items (not presented here), suggests that such an interaction was obtained for the focal brand when it was evaluated by itself.

## References

- Daryl J. Bem, "Self-Perception Theory," in L. Berkowitz (ed.), Advances in Experimental Social Psychology (New York: Academic Press, 1972).
- Bobby J. Calder and Barry M. Staw, "Self-Perception of Intrinsic and Extrinsic Motivation," Journal of Personality and Social Psychology 31(1975), 599-605.
- Ruby R. Dholakia and Brian Sternthal, "Highly Credible Sources: Persuasive Facilitators or Persuasive Liabilities?" Journal of Consumer Research 3(1977), 223-232.
- Jonathan L. Freedman and Scott C. Fraser, "Compliance Without Pressure: The Foot-in-the-Door Technique," Journal of Personality and Social Psychology 4(1966), 195-202.
- Mark 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(1973), 129-137.
- John B. Pryor and Mitchel Kriss, "The Cognitive Dynamics of Salience in the Attribution Process," Journal of Personality and Social Psychology 35(1977), 49-55.
- Peter Reingen and Jerry Kernan, "Compliance with an Interview Request: A Foot-in-the-Door, Self-Perception Interpretation," Journal of Marketing Research, 14(1977), 365-369.
- Carol A. Scott, "The Effect of Trial and Incentives on Repeat Purchase Behavior," Journal of Marketing Research, 13(1976), 263-269.
- Carol A. Scott, "Modifying Socially-Conscious Behavior: The Foot-in-the-Door Technique," Journal of Consumer Research, 4(1977), 156-164.
- Carol A. Scott and Richard F. Yalch, "A Test of Self-Perception Explanation of the Effects of Rewards on Intrinsic Interest," Journal of Experimental Social Psychology, 14(1978), 180-192.
- D. Sears and J. L. Freedman, "Selective Exposure to Information: A Critical Review," Public Opinion Quarterly, 31(1967), 194-213.
- Mark Snyder and M. R. Cunningham, "To Comply or Not Comply: A Testing of the Self-Perception Explanation of the Foot-in-the-Door Technique," Journal of Personality and Social Psychology, 31(1975), 64-67.
- William R. Swinyard and Michael R. Ray, "Advertising-Selling Interactions: An Attribution Theory Experiment," Journal of Marketing Research, 14(1977), 509-516.
- Alice M. Tybout, "The Relative Effectiveness of Three Behavioral Influence Strategies as Supplements to Persuasion," Journal of Marketing Research, 15(1978), 229-242.
- S. Valins, "Persistent Effects of Information About Internal Reactions: Ineffectiveness of Debriefing," in H. London and R. E. Nisbett (eds.), Thought and Feeling: Cognitive Alteration of Feeling States (Chicago: Aldine, 1974).
- B. Winer, Statistical Principles in Experimental Design (New York: McGraw Hill, 1971).



THE ROLE OF SEXUALLY-ORIENTED STIMULI  
IN ADVERTISING: THEORY AND LITERATURE REVIEW

R. Dale Wilson, The Pennsylvania State University  
Noreen K. Moore (student), The Pennsylvania State University

Abstract

In recent years there has been an increase in the use of sexual themes and erotic stimuli in advertising. This paper proposes some theoretical concepts through which the effects of these stimuli may be understood. Empirical research from consumer behavior and psychology is reviewed. Together, the theoretical and empirical components of the paper provide guidance for future research.

Introduction

It almost goes without saying that today's consumers are exposed to more sex in advertising than ever before. This trend towards the increasing depiction of sexually-oriented themes in advertising is evident upon even the most casual inspection of print and broadcast media. In general, advertising copywriters use nudity, romantic themes, or suggestiveness to draw attention to specific advertisements and their sponsoring products. In the marketplace of today, the use of such explicitly or implicitly sexually-oriented ads seems to be especially popular for parity products that must compete intensely for consumers' attention (Danielenko, 1974).

The use of sex in advertising has important social, managerial, and public policy implications. These implications hinge, of course, on the desired portrayal of women in contemporary society, the conditions under which advertising and marketing managers should and should not use sexually-oriented stimuli in communicating with consumers vis-a-vis competitors, and the impact of such communications on certain consumer groups (e.g., children). Other broad issues, such as the overall image of corporate advertising and consumers' rights in avoiding media content which they deem offensive also must be considered when sexually-oriented advertising is evaluated.

But despite the importance of sexually-oriented advertising and its increased use, little research has been directed towards measuring the impact of such advertisements on consumer awareness, attitudes, or behavior. And, perhaps more importantly, little thought seems to have been given to those theoretical concepts which may be useful in understanding the effectiveness of sexually-oriented ads. Most of the available information takes the "Look Ma--No Theory" approach that has recently been discussed by Jacoby (1976, 1978). By and large, there has been no effort to rely on theory to suggest those variables which are most important in understanding consumer response to sexually-oriented advertising stimuli; and the differences in individual psychological make-up which can cause opposite consumer reactions to the same ad are not well delineated at the current time. As a result of these shortcomings, very little is known about sex in advertising and how it may be expected to interact with other variables to arouse, motivate, offend, or otherwise affect target and non-target consumer populations.

The purpose of this paper is to help fill the void caused by a lack of theoretical and empirical attention to the area of sexually-oriented advertising communications. First, the paper reviews several potentially useful theoretical bases for predicting and understanding the effect of sexy advertisements. In this discussion, emphasis is placed on 1) the hypothesized differential effects of a

sexually-oriented advertisement versus a non-sexual theme and 2) why these differences might be expected to occur. Secondly, this article summarizes the limited amount of research from marketing and consumer behavior as well as from clinical and social psychology which uses sexually-oriented objects as stimuli. Lastly, the paper discusses some important directions for future research.

Theoretical Considerations

Arousal

An understanding of the concept of arousal is central to the understanding of the effects of sexually-based stimuli on consumer attention and recall. Arousal does not seem to be well understood in a consumer context, and little consumer research has been directed towards uncovering its specific role in the processing of brand-related information by consumers. Its importance, however, has been documented by direct attention in at least two comprehensive models of consumer behavior (Hansen, 1972; Howard, 1977) and in the area of consumer novelty seeking behavior (Venkatesan, 1973).

Arousal, usually defined as the degree of tension in the body, is a physiological state which gives rise to attention and search in the consumer decision-making process (Howard, 1977, p. 136). In the context of sexually-oriented stimuli, arousal caused by such stimuli can be thought to relate directly to attention toward the particular advertisement and may be mediated by the consumers' innate motivational state. Sexually-oriented ads may also relate to consumer reaction through stimulus ambiguity since an unexpected or surprise stimulus (one form of ambiguity as defined by Howard and Sheth (1969, p. 158) and Howard (1977, p. 140) guides the consumer search process. In this way, arousal can be thought to be a form of consumer motivation though its tension-producing effects are not well delineated.

The impact of environmental stimulation by a sexually-oriented advertisement may be illustrated by an "equilibrium" model in which a consumer is initially at an equilibrium with regard to sex. When a sensory input (an erotic ad, for example) is presented to the consumer, a disequilibrium state may be created since the incoming cue is associated with an innate motive (sexual activity). Psychological and perhaps physiological tension produced by the disequilibrium may then cause increased cognitive activity directed towards the ad and/or the advertised product. This enhanced level of information processing then will interact with the consumer's stable value system to produce an affective evaluation of the ad and, consequently, the product being advertised. The overall product image can thus be affected by the interaction of the type of stimulus object (and the context in which the stimulus is presented) with the structure of values and beliefs held by the receiver of the advertising message.

Of major importance in this process is the increased cognitive activity produced by the disequilibrium amount of arousal. A non-sexually-oriented stimulus may not have caused the same degree of arousal. The primary conclusion that must be drawn from arousal theory is that, relative to a non-sexy ad, a sexually-oriented communication will (perhaps significantly) increase the level of consumer

attention to the ad. However, there is no reason to believe that increased attention will positively affect consumer attitudes and purchase behavior. On the contrary, it is very likely that for consumer groups with certain sets of values, the image of the brand may deteriorate.

In addition to its attention to individual consumer differences in the form of values, arousal theory can also be used to understand the impact of sexual stimuli deemed irrelevant in certain advertising situations. Probably the best known form of irrelevant sex in advertising is the inclusion of a female model with no logical relation to the product, who performs only a decorative function (cf., Courtney and Lockeretz, 1971; Venkatesan and Losco, 1975). Available theory and research suggests that advertisements with such "sex object" cues may be tuned out since arousal actually inhibits the impact of irrelevant information (Howard, 1977, p. 143). Especially in a routinized (or nearly routinized) buying decision, such an advertising communication would not indicate that the advertised product would meet the consumer's rather well-defined set of choice criteria. Therefore, the effectiveness of the ad, both in terms of cognitive affect and search motivation, may be much less than for an ad which depicts the brand in a more meaningful, product-relevant way.

Arousal theory also suggests some considerations regarding the measurement of sensory response to sexually-oriented advertising. Relative to an effective non-sexual ad, an effective sexy advertisement might be expected to cause physiological responses that could be tracked by monitoring devices. Among these are galvanic skin response, pupil dilation, eye movements, response speed, heart beat, and breathing patterns. Recall and attitude measurement could determine consumers' evaluative responses and intended behavior, as well as the amount of learning which occurred during the exposure to the sexually-oriented ad.

#### Selective Perception

Another of the set of theoretical concepts which may be used to understand the effects of sexually-oriented advertising stimuli is selective perception. Usually defined as the complex process by which consumers select, organize, and interpret sensory stimulation (Berelson and Steiner, 1967, p. 141), selective perception can be used as the theoretical basis for studying the strength and ambiguity of sexually-oriented ads. Research results in the area of absolute thresholds, differential thresholds, and sensory projection of ambiguous stimuli may provide important clues as to how consumers ultimately interpret and evaluate advertisements using sexual themes.

Often used to explain consumer response to pricing stimuli (cf., Monroe, 1973), the absolute threshold marks the lower limit of sensory stimulation. In the case of sexually-oriented advertising, such an ad must be of a certain intensity before its sexual meaning may be grasped. "Suggestive" messages must, therefore, fall above the absolute perceptual threshold before their sexual connotations are understood by those who see or hear the ad.<sup>1</sup> Individual differences and environmental factors will also affect the consumer's ability to perceive sexually-implicit stimuli.

Also useful in understanding perceptual influences on advertising stimuli is the differential threshold, the minimum difference in two or more stimuli which causes a

<sup>1</sup>An excellent illustration of the role of absolute threshold levels in explaining consumer response to sexual innuendo in advertising is provided by Johnson and Satow (1978) in their discussion of the "Edith Bunker reaction" to a BIC Razor commercial by a group of older women.

change in discrimination among those stimuli. This minimum amount of change necessary to produce a "just noticeable difference" (or j.n.d.) may be useful in explaining the difference between a sexually-oriented stimulus perceived as being tasteful, meaningful, and appealing and another stimulus perceived as tasteless and offensive. In other words, the degree of intensity of the stimulus or the context in which the stimulus is presented will make the difference in consumers' perception of and reaction to sex in advertising.

Closely related to threshold levels is the ambiguity of the stimulus object. Theoretical discussions of stimulus ambiguity (e.g., Berelson and Steiner, 1967, pp. 155-57; Howard and Sheth, 1969, pp. 156-67) stress that an ambiguous object is likely to be perceived in a way which is consistent with the individual's experiences, prior beliefs, and enduring values. In other words, "perceptual constancy" causes consumers to perceive ambiguous stimuli in congruence with their existing cognitive structures. Further, there seems to be an observable relationship between stimulus ambiguity and arousal which affects an individual's response to sexually-oriented stimuli. In an applications sense, advertisers may be well advised to employ moderately ambiguous themes which may be perceived as sexually-oriented by some viewers (and thus cause arousal) but may fall below the threshold of other viewers (and thus may avoid offending them). As noted previously, however, there may be a significant difference between arousal caused by observing a sexy ad and enhanced brand image or persuasion due to the ad.

#### Self-Concept

Notwithstanding the importance of arousal and selective perception in understanding the impact of sexually-oriented communications, self-concept (or self-image) theory may also make a contribution. The reader is referred to discussions of the theoretical and empirical work in self-concept in Hansen (1972, pp. 362-3, 377-80), Howard (1977, pp. 89-92), and Engel et al. (1978, pp. 141-3) for a thorough review of this topic.

In the context of sexually-oriented advertising, it seems important to consider the influence of both actual and ideal self-concept. Anecdotal evidence to this effect is provided in Johnson and Satow's (1978) discussion of female fantasy ads. Theoretically, a viewer of a sexually-oriented advertisement would be expected to evaluate the ad on the basis of the role of the actors or models in the ad. In the case of a "sexy" female model, a female viewer would be more likely to project herself into the situation if she perceives herself as actually being sexy (actual self-concept) or, perhaps more importantly, if she wants to be sexy (ideal self-concept). The crucial point is that the stimulus should be congruent with either actual or ideal self-image in order to serve as a link between the viewer of the ad and the advertised product.

The role of self-concept may also be valuable in explaining the negative reaction of feminist women's groups to sex in advertising, especially the use of female models solely as sex objects. These critics would not tend to project themselves into the advertisement since their values are diametrically opposed to those embodied in the female model. The result of this situation is an appropriate negative response. The advertiser is perceived as attempting to manipulate consumers by using a socially undesirable appeal (degradation). Other female viewers, who are not cognizant of a manipulative intent and whose values are consistent with the model's role, may respond favorably to the same ad. Self-concept theory suggests that a sexually-oriented stimulus must be appropriate for the product and congruent to the values of the recipient of the ad in order for the desired response or projection to occur. Only then will a favorable brand image be created or maintained by association with the stimulus.

## Distraction

Originally stemming from the early research in communications conducted in conjunction with the Yale Communication and Attitude Change Program, the study of distraction effects now plays a central role in understanding the impact of persuasive messages. While the experimental evidence seems contradictory (Karlins and Abelson, 1970, pp. 15-18), some research results seem to have strong implications for the role of sex in advertising. Based upon the findings reported by Haaland and Venkatesan (1968) and Zimbardo et al. (1968) and others, communications employing moderately distracting stimuli which positively reinforce the consumer will probably be more persuasive than other types of distraction and perhaps than no distraction at all.

In terms of sexually-oriented advertising, the available research would seem to indicate that sexual distraction, even if it is irrelevant to the product being advertised, can play a positive role in the communications process. Even though distraction hypotheses do not form the theoretical bases for the reported research on sex in advertising, there is evidence that distraction provides the rationale for its practical use. Danielenko (1974) cites an argument for using sex in advertising based on the premise of blocking counterarguments of a hostile receiver. The claim is that a sexy stimulus distracts the consumer and no counterarguments are initiated. The result is that the consumer remembers less about the ad, but is persuaded more than if a non-distracting stimulus is used.

With respect to this interpretation of the Yale studies in an advertising context, one qualifier seems appropriate. In advertising a sexual stimulus is likely to be the attention getter; successful processing of the ad content then depends on the transfer of attention to the advertising message. The chance of this happening may decrease as the irrelevancy of the stimulus to the product increases. The success of the irrelevant stimulus in the Yale may stem from the fact that the primary focus of was the persuasive message, so that some processing of that message was guaranteed. This could be investigated empirically; in any case distraction theory seems a useful foundation for further research.

## Aggression Theory

The long-held contention that sexual and aggressive motives are correlated provides the structure for much of the current literature in aggression theory. Space limitations do not allow a review of this literature, but the interested reader is referred to Bandura (1971), Baron (1974), Jaffe et al. (1974), and Baron (1978) for current perspectives. The important point from this line of thinking is the possibility that there may be a connection between sexually-oriented stimuli and aggressive behavior on the part of individuals who are exposed to these stimuli.

A thorough review of the aggression/sexual stimuli literature reveals no managerial significance coming from this body of knowledge. But, undoubtedly, there would be strong social and public policy implications if the data revealed evidence of a positive relationship (or, worse yet, a cause and effect relationship) between viewing sexually-oriented stimuli and aggressive behavior. This possibility serves as the reasoning behind the inclusion of a brief review of aggression theory results. These results, as well as results from other areas relevant to sexually-oriented advertising, are summarized in the following section of this paper.

### Evidence from Empirical Research

The frequent use of sexual stimuli in advertising testi-

fies to a widespread belief in its effectiveness. However, little research has been directed at justifying this faith or delineating the nature of the presumed benefits. This section reviews six marketing studies that have addressed this topic and summarizes some psychological literature on human response to erotic stimuli.

The psychology literature can be divided into two areas which seem relevant to advertising: (1) sex differences in response to erotica and (2) the effects of erotic stimuli on aggressive tendencies. One important limitation on the generalizability of these results to advertising concerns the degree of eroticism of the stimuli used. Psychological research typically contains pictures or films of explicit sexual acts which by comparison make the most daring advertising erotica seem mild. Responses to such disparate levels of sexual stimuli may differ in kind as well as degree. Despite this limitation, this research may provide insight into the effects of sexually-oriented communications messages.

## Aggression

An issue of great social concern is how sexually-arousing stimuli affects aggressive behavior. Aside from the obvious implications for censorship of pornography there is an ethical problem regarding the use of erotic advertisements in marketing. Many psychology experiments in the last decade have attempted to resolve this issue. Results of these studies have varied.

Contrary to prior expectations, mildly erotic stimuli most frequently decreased subsequent aggressive behavior by subjects angered by a confederate (Baron, 1978; Frodi, 1977; Baron and Bell, 1977; Donnerstein et al., 1975; Baron, 1974). Exposure to highly explicit stimuli increased aggression in angered subjects in some instances (Donnerstein and Barrett, 1978; Jaffe et al., 1974; Meyer, 1962), but had no effect in others (Baron and Bell, 1977; Donnerstein et al., 1975). Studies which manipulated the gender of the confederate found no variations in aggression due to this factor (Donnerstein and Barrett, 1978; Jaffe et al., 1974).

Compared to the stimuli used in these psychological studies, sexually-oriented stimuli generally used in legitimate advertising are not likely to be classified as highly erotic. At the current time, the available studies provide no evidence that the use of mildly erotic stimuli in advertising is socially irresponsible due to its provoking aggressive behavior.

## Sex Differences in Response to Erotica

Psychological research has recently challenged a time-honored stereotype about gender differences in response to erotic stimuli. In 1953, a survey of American women reported that they were not aroused by erotic stimuli (Kinsey et al., 1953), and subsequent surveys periodically corroborated this finding. Contradictions began to emerge only in the experimental research of the 1970's. Discussion of these works will cover two primary areas: sex differences in arousal by erotica, and sex differences in affective/evaluative ratings of erotica. As discussed previously, sexual arousal may have an effect on information seeking or serve as a distractor from counterarguing, thus increasing advertising effectiveness. In addition, the affect and evaluation of an advertising stimulus could conceivably generalize to the advertisement, product, and sponsoring company. Differential response by males and females to erotic stimuli would imply that effective use of this device is dependent on the ad's target audience.

In several studies, measures of general sexual arousal when subjects were shown erotica revealed no sex differences (Schmidt, 1975; Schmidt and Sigusch, 1970). Males reported a greater overall arousal in research by Sigusch et al.

(1970), yet there was no sex difference in frequency of genital excitation in response to those stimuli. Empirical evidence also contradicted the stereotype of women being more aroused by romantic, emotional stimuli (Fisher and Byrne, 1978; Schmidt, 1975).

Despite these findings, there do still appear to be some sex differences in reaction to erotica. In Schmidt's (1975) study, women reported more arousal than men to a "same sex" stimulus (referring to the sex of the subject observer and the sex of the actor in the stimulus). And in other research (Griffitt, 1973; Schmidt and Sigusch, 1970), females reported less arousal than men to the depiction of less conventional sexual acts. Women also reported more emotional arousal during and after observation of erotic stimuli (Sigusch and Schmidt, 1970). It must be emphasized that even when mean differences between sexes were significant, individual variation was great, with segments of the female samples reporting more arousal than the average male in several such instances (Schmidt and Sigusch, 1970; Schmidt et al., 1970).

The same studies also measured affect and evaluation for the erotic stimuli, and a pattern of sex differences did emerge. Females were more likely than males to label the stimuli as pornographic and give ratings of disgust, anger and nausea (Fisher and Byrne, 1978; Griffitt, 1970). Women also tended to evaluate stimuli using "same sex" and romantic stimuli more favorably and erotica depicting less conventional sexual acts less favorably than did males (Schmidt, 1975; Sigusch et al., 1970; Schmidt and Sigusch, 1970). An exception to this pattern was reported by Fisher and Byrne (1978), who found no sex differences in affect for stimuli, regardless of erotic theme.

From the evidence cited thus far, male and female subjects seem to have similar physical reactions to sexual stimuli. Both respond to erotica with arousal, and both generally respond more strongly to stimuli involving the opposite sex. The cognitions which accompany that arousal appear to vary more. Females have a greater tendency to evaluate erotic stimuli negatively, with possible adverse implications for product advertising. The more favorable evaluation of "same sex" stimuli by women would indicate that their use would be more appropriate for female than male products. However, there may be a greater tendency in an advertising situation to project oneself into the illustration rather than regarding the actor in the stimulus as an object of desire, thereby mitigating the negative reaction by males to same sex stimuli.

#### Consumer Research on Sexual Stimuli in Advertising

In the marketing literature, six studies have examined the use of sexual stimuli in advertising. This small body of work is, unfortunately, not cohesive. The studies have approached different aspects of the phenomenon, using widely varying designs and variables. A comparison of the major characteristics of each is given in Table 1. Due to their disparity, each study can best be analyzed in the context of the major question it addresses.

#### Consumers' Attitudes on the Amount of Sex in Advertising.

Wise, King, and Merenski (1974) asked a large sample of college-age respondents and their parents to rate the extent of their agreement with the statement "Advertisers make too much use of sexual appeals in ads." The only significant variables were age and sex; females and older respondents agreed more frequently. The finding that females have a more negative affect for erotica than do males is consistent with psychological research.

Application of these results to a decision about whether to use a sexual appeal for an older and/or female target market is nonetheless unwarranted. The attitude statement used is broadly phrased. Agreement may indicate that only some sexual appeals are sexist or in bad taste, that all

erotic advertisements are immoral, or even that some reach an inappropriate audience such as children. Reactions of females and older consumers to a tasteful, appropriate sexual appeal could well be favorable. The dimensions of the sexy ads to which these groups objected should be explored.

#### Dimensions Which Consumers Perceived in Sexual Stimuli.

Morrison and Sherman (1972) asked Ss to rate a total of 100 advertisements on six dimensions. Clustering techniques were used to determine similarities across respondents' ratings for both sexes. The research pointed to several interesting findings. First, women differed most in their tendency to perceive nudity; of those who did perceive it, 68% reported being aroused. Secondly, not all men perceived nudity in the ads. Sexual arousal was not contingent on the degree of nudity in the ad; like women, men can be aroused by suggestive and mildly sexually-oriented stimuli. Thirdly, women tended to report both romantic theme and sexual arousal perceptions, or neither. This contradicts the findings from psychological research that arousal for women is not contingent on a romantic, affectionate theme. Lastly, female groups did not differ in their tendency to perceive copy suggestiveness, while male groups did. Copy suggestiveness may, therefore, be a more important variable for women than formerly thought.

A few methodological points in the Morrison and Sherman study prevent unqualified acceptance of these results. The input for the clustering involved ratings on different ads. Clustering respondents on the similarity of their perceptions of different stimuli may have resulted in the grouping of individuals who had seen similar ads, rather than the grouping of people who have a general tendency to perceive similar dimensions.

Sexual Stimuli and Recall/Recognition Measures. In a study designed to investigate brand recall, Steadman (1969) found that non-sexual illustrations resulted in better recall of brand names with which they were paired than did sexual illustrations. This effect remained in a retest one week later. Recall did not vary according to the degree of eroticism of the sexual illustrations. Respondents who favored use of sex in advertising had better recall of brand names paired with sexual illustrations than did other respondents. Although plagued by problems of external validity, the findings of Steadman's research have found further empirical support.

Alexander and Judd (1978) partially replicated the Steadman study and included several methodological improvements. Their study found poorer "brand recall" for those ads with a female model as opposed to ads which include inanimate objects. Surprisingly, there was no difference in recall for brands paired with a picture of a model's face and neck and brands paired with a full length picture of a nude model. This fact suggests the possibility that the human element rather than nudity per se is responsible for the worse recall. The higher interest created by the human pictures may have distracted subjects from learning "brand names." However, that same feature could be advantageous in a realistic setting, since it may increase the likelihood of the ad getting any attention in the first place.

The most convincing evidence for the lack of impact of sexual stimuli on brand name recognition is presented by Chestnut et al. (1977). Here, the presence of decorative models improved ad recognition but had no effect on brand name recognition. Overall, ad recognition was significantly greater than brand name recognition. These differences were attributed to variations in the encoding process.

The existence of a badly needed theoretical framework in the Chestnut et al. study generates some provocative questions for future research. Interesting issues include the roles of repetition and stimulus/product congruency in the encoding process. It seems clear that an approach based

TABLE 1

## CHARACTERISTICS OF MARKETING RESEARCH ON SEXUAL STIMULI IN ADVERTISING

	Steadman (1969)	Morrison and Sherman (1972)	Wise, King, and Merenski (1974)	Peterson and Kerin (1977)	Chestnut, LaChance, and Lubitz (1977)	Alexander and Judd (1978)
Product Used	None	16 (Unspecified)	N.A. <sup>a</sup>	Wrench Set; Body Oil	50 (Unspecified)	15 (Unspecified)
Predictor Variables	Erotic Content; Time Lag; Attitude Toward Use of Sexy Ads	Product Sexual Connotation; Degree of Nudity; Realism; Romantic Content; Suggestive Copy; Sexual Arousal	Age; Sex; Race; Income; Edu- cation; Major Curricu- lum	Product Type; Respondent Sex; Model Nudity	Presence of Model; Type of Recog- nition; Test (Entire Ad or Brand Name Only)	Model Nudity; Attitudes Toward Sexual Stimuli; Demographic Variables
Criterion Variables	"Brand Name" Recall	Membership in Cluster Groups	Attitude Toward Use of Sex Appeals in Advertising	Product Quality; Company Reputation; Ad Appeal	Ad Recogni- tion Brand Name Recognition	"Brand Name" Recall
Blocking Variables	None	Sexual Arousal; Product Class	None	None	None	None
Media	Print	Print	N.A.	Print	Print (On Slides)	Print (On Slides)
Realistic Ads	No	Yes	N.A.	Yes	Yes	Yes
Ads in Natural Setting	No	No	N.A.	No	No	No
Total Number of Stimuli	12	100	N.A.	8	200	15
Number of Stimuli Per Respondent	12	10	N.A.	1	50	15
Number of Presentations	Unknown (24 Hour Exposure)	1	N.A.	1	1	2
Experimental Design	Mixed Design; One Between Two Within Ss Factors	No	No	2x2x4 Factorial	2x2 Factorial	Mixed Design; One Between One Within Ss Factor
Analytical Technique	Chi-Square Statistic	Q-Analysis and Heuristic Clustering Analysis	Chi-Square Statistic	ANOVA	ANOVA	ANOVA; Scheefé Contrasts; F-Tests
Sample Size and Composition	29 Business Post- Graduates; 31 Business- men	100 Male; 100 Female Undergradu- ates	621 College Age Respondents and 589 of Their Parents	112 Male; 112 Female South- west Metro- politan Adults	103 Male Undergradu- ates	181 Uni- versity Students
Major Design Problems	Questionable	Questionable	No	No	No	No

<sup>a</sup>Not Applicable

on these theoretical perspectives holds the promise of obtaining less equivocal results than have been thus far demonstrated.

Sexual Stimuli, Ad Appeal, Product Quality, and Company Reputation. Peterson and Kerin (1977) conducted an experiment in which "mock up" ads were varied by product (body

oil vs. wrenches) and model (none vs. demure vs. seductive vs. nude). The seductive model/body oil ad elicited the best evaluations on ad appeal, product quality, and company reputation. Nude model/body oil received the worst evaluation on these dimensions. Males tended to rate all of the ads more favorably on each variable than did females.

It was expected that favorability of evaluations would increase as product/model congruency increased, which was supported by the overall worse evaluation of the wrench ads which included a model. But some exceptions are puzzling. The nude model/wrench ad was evaluated more favorably on appeal and product quality than was the nude model/body oil ad, even though the latter is undoubtedly more congruent.

The overall relatively negative evaluations for the nude model ads should not necessarily exclude their use. The study did demonstrate the effect of sexual appeal/product congruency; perhaps the discovery of other influential variables will reveal under what conditions nude stimuli may profitably be used. The Peterson and Kerin (1977) study is only a beginning in uncovering such variables and interactions, but it demonstrates a successful use of the experimental paradigm necessary to do so.

#### Research Recommendations

The research reviewed in this paper has identified many factors which may influence the effectiveness of sexually-oriented advertising communications. Determining the relative importance of each and the ways in which they interact is a major undertaking for future research. Focus groups, like those used by Johnson and Satow (1978), may be helpful in identifying and probing these dimensions. The need for experimentation and information processing research methods seems an inescapable conclusion.

Multiple measures of advertising effectiveness seem essential for future research, but design should vary somewhat with the measures of concern. For instance, a realistic setting is of great importance for studies of brand recall, but may be of little consequence in measuring arousal. New physiological measurement techniques might be borrowed from psychological research on erotica. For example, reliability and validity checks on a measurement device for female sexual arousal called a photoplethysmographic probe have shown excellent results. The measure appears to be without artifacts and does not seem to interfere with attention to the stimuli. Both of these characteristics are improvements over a pencil and paper measure of arousal (Wincze, Hoon, and Hoon, 1977).

One problem inherent in any experimental work dealing with advertising effectiveness is the quality of the experimental ads. An effective advertisement has multiple components, and neglecting any of these threatens to negate the managerial value of research findings. For instance, the lighting, facial expression, body position, background, etc, that most enhance a message accompanied by a nude model are likely to differ from those most appropriate when no model or a clothed model is present. Care must be taken to make realistic, high quality ads without introducing confounding factors. Realization of this goal will be difficult, but it merits the additional effort.

#### Conclusions

The purpose of this paper has been to suggest possible theoretical approaches to the study of sexual stimuli in advertising and to summarize the relevant literature. Several conclusions are forthcoming from this review. First, the role of sexually-oriented advertising stimuli may be more clearly understood through the theoretical constructs discussed in this paper. It is also evident that the operation of each of these concepts is mediated by the consumer's enduring value system. The understanding of sex in competitive advertising is therefore contingent upon the consideration of personal values in the context of such theoretical constructs.

Another major conclusion stemming from this paper is the apparent opportunity for research in this area. Such research would offer practical input into the questions of advertising's role in society and the managerial uses of sexually-oriented advertising. It is imperative that any such empirical work employ stimuli designed to be realistic representations of advertising. Research that is tied to a theoretical framework and possesses adequate external validity can develop the promising area of sexually-oriented stimuli into an important aspect of communications research.

#### References

- M. Alexander and B. Judd, "Do Nudes in Advertisements Enhance Brand Recall?", Journal of Advertising Research, 18(February, 1978), 47-50.
- A. Bandura, "Social Learning Theory of Aggression," in Control of Aggression: Implications from Basic Research, ed. J. F. Knutsen, (Chicago: Aldine, Atherton, 1974).
- R. A. Baron, "The Aggression-Inhibiting Influence of Heightened Sexual Arousal," Journal of Personality and Social Psychology, 30(September, 1974), 318-22.
- , "Aggression-Inhibiting Influence of Sexual Humor," Journal of Personality and Social Psychology, 36(February, 1978), 189-97.
- and P. A. Bell, "Sexual Arousal and Aggression by Males: Effects of Types of Erotic Stimuli and Prior Provocation," Journal of Personality and Social Psychology, 35(February, 1977), 79-87.
- B. Berelson and G. A. Steiner, Human Behavior: Shorter Edition (New York: Harcourt, Brace and World, 1967).
- D. Byrne, J. D. Fisher, J. Lamberth and H. E. Mitchell, "Evaluations of Erotica: Fact or Feelings?", Journal of Personality and Social Psychology, 29(January, 1974), 111-16.
- R. W. Chestnut, C. C. LaChance and A. Lubitz, "The 'Decorative' Female Model: Sexual Stimuli and the Recognition of Advertisements," Journal of Advertising, 6(Fall, 1977), 11-14.
- A. E. Courtney and S. W. Lockeretz, "A Women's Place: An Analysis of the Roles Portrayed by Women in Magazine Advertisements," Journal of Marketing Research, 8(February, 1971), 92-5.
- R. Danielenko, "Do Sexy Ads Sell Products?", Product Management, 40(February, 1974), 21-6.
- E. Donnerstein and G. Barrett, "Effects of Erotic Stimuli on Male Aggression Toward Females," Journal of Personality and Social Psychology, 36(February, 1978), 180-88.
- , M. Donnerstein and R. Evans, "Erotic Stimuli and Aggression: Facilitation or Inhibition," Journal of Personality and Social Psychology, 32(August, 1975), 237-44.
- J. F. Engel, R. D. Blackwell and D. T. Kollat, Consumer Behavior, 3rd Ed. (Hinsdale, Illinois: Dryden Press, 1978).
- W. A. Fisher and D. Byrne, "Sex Differences in Response to Erotica? Love vs. Lust," Journal of Personality and Social Psychology, 36(February, 1978), 117-25.
- A. Frodi, "Sexual Arousal, Situational Explicitness, and Aggressive Behavior," Journal of Research in Personality, 11(March, 1977), 48-58.

- W. Griffitt, "Response to Erotica and Projection of Response to Erotica in the Opposite Sex," Journal of Experimental Research in Personality, 6(April, 1973), 330-8.
- G. Haaland and M. Venkatesan, "Resistance of Persuasive Communications: An Examination of the Distraction Hypothesis," Journal of Personality and Social Psychology, 9(June, 1968), 167-70.
- F. Hansen, Consumer Choice Behavior: A Cognitive Theory (New York: Free Press, 1972).
- J. A. Howard, Consumer Behavior: Application of Theory (New York: McGraw-Hill, 1977).
- J. A. Howard and J. N. Sheth, The Theory of Buyer Behavior (New York: John Wiley and Sons, 1969).
- J. Jacoby, "Consumer Research: Telling It Like It Is," in Advances in Consumer Research, Vol. 3, ed., B. B. Anderson (Cincinnati: Association for Consumer Research, 1976), 1-11.
- J. Jacoby, "Consumer Research: A State of the Art Reviews" Journal of Marketing, 42(April, 1978), 87-96.
- Y. Jaffe, N. Malamuth, J. Feingold and S. Feshback, "Sexual Arousal and Behavioral Aggression," Journal of Personality and Social Psychology, 36(February, 1974), 189-97.
- D. K. Johnson and K. Satow, "Consumers' Reactions to Sex in TV Commercials," in Advances in Consumer Research, Vol. 5, ed., H. K. Hunt (Chicago: Association for Consumer Research, 1978), 411-14.
- M. Karlins and H. I. Abelson, Persuasion: How Opinions and Attitudes are Changed, 2nd Ed., (New York: Springer, 1970).
- A. C. Kinsey, W. B. Pomeroy, C. E. Martin and P. H. Gebhard, Sexual Behavior in the Human Female (Philadelphia: Saunders, 1953).
- J. Mann, L. Berkowitz, J. Sidman, S. Starr and S. West, "Satiation of the Transient Stimulating Effect of Erotic Films," Journal of Personality and Social Psychology, 30 (December, 1974), 729-35.
- T. Meyer, "The Effects of Sexually Arousing and Violent Films on Aggressive Behavior," Journal of Sex Research, 8(November, 1962), 324-31.
- K. B. Monroe, "Buyers' Subjective Perceptions of Price," Journal of Marketing Research, 10(February, 1973), 70-80.
- B. J. Morrison and R. C. Sherman, "Who Responds to Sex in Advertising?," Journal of Advertising Research, 12(April, 1972), 15-19.
- R. A. Peterson and R. A. Kerin, "The Female Role in Advertising: Some Experimental Evidence," Journal of Marketing, 41(October, 1977), 59-63.
- G. Schmidt, "Male-Female Differences in Sexual Arousal and Behavior During and After Exposure to Sexually Explicit Stimuli," Archives of Sexual Behavior, 4(1975), 353-64.
- and V. Sigusch, "Sex Differences in Response to Psychosexual Stimulation by Films and Slides," Journal of Sex Research, 6(November, 1970), 268-83.
- V. Sigusch, G. Schmidt, A. Reinfeld and I. Wiedemann-Sutor, "Psychosexual Stimulation: Sex Differences," Journal of Sex Research, 6(February, 1970), 10-24.
- M. Steadman, "How Sexy Illustrations Affect Brand Recall," Journal of Advertising Research, 9(February, 1969), 15-19.
- M. Venkatesan, "Cognitive Consistency and Novelty Seeking," in Consumer Behavior: Theoretical Sources, ed., S. Ward and T. S. Robertson (Englewood Cliffs, New Jersey: Prentice Hall, 1973).
- and J. Losco, "Women in Magazine Ads: 1959-71," Journal of Advertising Research, 15(October, 1975), 49-54.
- J. P. Wincze, P. Hoon and E. F. Hoon, "Sexual Arousal in Women: A Comparison of Cognitive and Physiological Responses by Continuous Measurement," Archives of Sexual Behavior, 6(1977), 121-33.
- G. L. Wise, A. L. King and J. P. Merenski, "Reactions to Sexy Ads Vary With Age," Journal of Advertising Research, 14(August, 1974), 11-16.
- P. Zimbardo, E. Ebbesen and S. Fraser, "Emotional Persuasion: Arousal State as a Distractor," unpublished manuscript, Stanford University, 1968.

DEMOGRAPHIC AND COGNITIVE FACTORS  
INFLUENCING VIEWERS EVALUATIONS OF "SEXY" ADVERTISEMENTS

Donald Sciglimpaglia, San Diego State University  
Michael A. Belch, San Diego State University  
Richard F. Cain, Jr. (student) San Diego State University

Abstract

A trend toward the increasing use of sex in advertisements can hardly be disputed. One need only browse through the newspaper or magazines or tune in the television to quickly become aware of the extensive use of some form of sexual content or appeal, whether blatant or suggestive. As advertisers continue to increase the frequency with which sex-involved strategies are employed, the need for evaluating the effect of such a strategy also becomes increasingly important.

A review of the literature regarding the use of sex in ads reveals conflicting results. Opposite opinions are presented in respect to consumers' attitudes toward the amount of sex presented (Ad Age, 1973, Ad Age, 1968, TV Guide, 1975) and their response to the same (Wise, King, Merenski, 1974). The results are more consistent with respect to the effects on brand name recognition and recall, as Chestnut, LaChance, and Lubitz (1977) and Steadman (1969) essentially agree that the use of female models is not likely to increase recognition for the brand. In respect to differences in perceptions, Morrison and Sherman (1972) found that females were more likely to report on the suggestiveness of ad copy than were males, while Wise, King and Merenski found that age of the viewer has an effect on the perceived usefulness of sex in advertising. Finally, Peterson and Kerin (1977) found that ads using female nudes were consistently perceived as less appealing and that those products associated with the nude ad were seen as of lower quality. Peterson's conclusions--when considered with those presented earlier--seem to indicate that the results of using sex in advertising is likely to vary according to who is viewing the ad; the nature of the product advertised; and/or specific situational variables present at the time of the viewing.

The purpose of the study to be reported here, is to attempt to isolate factors that may be critical in the valuation of the use and effectiveness of sex in advertising. This isolation is more likely to allow for causal attributions and to serve as guidelines for the marketer concerned with the impact upon his/her product or service.

Methodology

An experimental study was conducted in which subjects were exposed to randomly ordered "print ads" for which no products or brand names were displayed. Included were ads in which sexual content was systematically varied. Subjects were asked to rate the "ads" on four affective scales. In addition, subjects completed a self-administered instrument which measured various personal difference factors which the authors considered as possibly influencing the evaluations.

Operationalizing Sex in Ads

Most of the earlier studies examining the use of sex in advertising have examined the same from a role portrayal or nudity perspective. For this study, we have attempted to define sex according to two constructs: (1) nudity (cf. Peterson and Kerin (1977)), and (2) suggestiveness. In respect to the former, both male and female

subjects were portrayed at three levels of nudity (see Appendix for detailed description). Freud's (1958) definition of suggestion ("a command or piece of information that triggers or arouses an idea in a person's mind") was utilized, and operationalized with both male and female subjects appearing together. This was done to reduce the possibility of perceiver bias and to enhance the perception of suggestiveness. This operationalization resulted in the use of nine print advertisements as follows:

1. Male nude
2. Female nude
3. Male partially clothed
4. Female partially clothed
5. Male fully clothed
6. Female fully clothed
7. Male and female nude (suggestive)
8. Male and female partially clothed (suggestive)
9. Male and female fully clothed (suggestive)

For a more specific elaboration of the description of each ad, see Appendix 1.

In addition to the nine test ads, eight dummy ads were also included in the task. These latter ads were photos of innocuous scenes such as sunsets, children, etc. Actual ads appearing in various print media were employed, with the exception of the male nude and male-female suggestive formats for which no examples were available. In those instances, photographs from magazines were used and presented in ad form. In each ad no brand name appeared, and no product was represented, thus eliminating possible intervening variables. Each print ad was transformed into a slide and presented to the respondent in various randomized sequences, along with the eight dummy ads. The respondent was asked to rate the ads on four dimensions including: 1) in good taste, 2) appealing, 3) interesting, and 4) offensive, in respect to their use in advertising.

Respondents

Respondents included 142 students from the business school at a major university on the West Coast. Of these, 42 were female. Mean age, 24.

Instrument

The complete instrument used contained four sections. The first three sections were self-administered cognitive measures. The fourth dealt with the subjects' evaluations of the ad portrayals on the four dimensions cited above and used six-point, forced choice semantic differential scales. The three cognitive measures were utilized to evaluate subjects' posture toward sex (both with respect to the societal context and with respect to personal orientation), attitudes toward role portrayals in advertising, and personal orientation with respect to social roles. The first of these was a modification of Ewell's Inventory of Values (1954), a thirty-eight statement Likert-type measure of overall social values. Since many of these value statements relate to personal and societal orientations with respect to sex, this measure of social values was used to evaluate an individual's sexual values. A high score on this scale was used to approximate a conservative sexual orientation while a low score represented a more liberal posture. The second cognitive measure was a twelve-item Likert



scale used to evaluate an individual's attitudes toward role portrayals in advertising and was adapted from a scale used by Lundstrom and Sciglimpaglia (1977). These items (e.g., "Ads show men as they really are" and "Advertisements treat women mainly as sex objects") were used as an overall attitudinal measure of how advertising portrays both men and women. A high score on this scale represented a more critical position toward contemporary role portrayals. The last measure utilized was Arnott's Role Orientation Scale (1972). This ten-item Likert Scale (e.g., "A woman should expect just as much freedom of action as a man") was used as a measure of the individual's personal orientation with respect to "proper" social roles of men and women. A low score on this scale was interpreted as a very traditional stance with respect to social roles while a high score was inferred to represent a more modern or "liberated" position.

#### Data Analysis

Responses of male and female subjects were evaluated using t-tests of means. In addition, Pearson product moment correlations were computed between total scores on each of the three personal difference variables and the evaluation of the experimental ads.

#### Results

The evaluations of each of the seventeen ads were first compared between male and female subjects. The results of the eight dummy ads showed little differences between sex, as was expected, and are not reported here. For the nine test ads, however, a different picture was obvious. The mean responses for males and females on the four affective scales (appeal, offensiveness, taste, interest) for each of the nine test ads are shown in Table 1 along with the computed significance of the differences in mean evaluations. Although these four affective scales were found to be highly intercorrelated, they are all displayed since they are assumed to tap different imputed cognitive domains.

TABLE 1  
EVALUATION OF  
ADVERTISING PORTRAYALS BY SEX OF RESPONDENT

Ad Portrayal	Sex <sup>1</sup>	Appeal <sup>2</sup> x (Sig.)	Offense <sup>3</sup> x (Sig.)	Taste <sup>4</sup> x (Sig.)	Interest <sup>5</sup> x (Sig.)
Male FC	F	2.24(0.00)	1.48(0.01)	1.98(0.02)	2.57(0.00)
	M	3.31	1.88	2.42	3.69
Female FC	F	3.36(0.04)	1.28 N.S.	2.36 N.S.	3.26 N.S.
	M	2.74	1.41	2.25	3.10
Male-Fem. FC (sug)	F	3.64(0.02)	3.69(0.04)	4.52(0.03)	3.69(0.03)
	M	2.78	3.09	3.96	3.06
Male PC	F	3.29(0.00)	2.19 N.S.	3.37 N.S.	3.51(0.00)
	M	4.64	2.59	3.55	4.70
Female PC	F	3.36(0.00)	2.79 N.S.	3.62(0.04)	3.74(0.00)
	M	1.85	2.41	3.07	2.10
Male-Fem. PC (sug)	F	3.10 N.S.	3.07 N.S.	3.98 N.S.	3.24 N.S.
	M	2.97	2.81	3.71	2.99
Male Nude	F	2.55(0.00)	2.57(0.01)	3.62(0.00)	2.65(0.00)
	M	4.87	3.29	4.40	4.93
Fem. Nude	F	4.34(0.00)	3.68(0.00)	4.37(0.00)	4.58(0.00)
	M	2.53	2.82	3.58	2.90

Male-Fem. F 3.88(0.00) 3.60(0.01) 4.24(0.02) 3.71(0.00)  
Nude (sug) M 2.69 2.86 3.64 2.83

<sup>1</sup>(N<sub>f</sub> = 42 N<sub>m</sub> = 100)

<sup>2</sup>(1 = very appealing, 6 = very unappealing)

<sup>3</sup>(1 = not offensive, 6 = offensive)

<sup>4</sup>(1 = in good taste, 6 = in poor taste)

<sup>5</sup>(1 = personally interesting, 6 = personally uninteresting)

Inspection of Table 1 reveals two main results: 1) both nudity and suggestiveness are factors which appear to influence the evaluation of "sexiness" of advertising portrayals, and 2) the evaluation of sexual content is a function of the sex of the evaluator. With respect to the influence of both nudity and suggestiveness, these data imply that the evaluation of sexual content becomes generally less positive as nudity increases. But, the suggestiveness of the portrayal, regardless of degree of nudity plays an intervening role. The three "suggestive" ads are consistently rated by both male and female subjects as being in poor taste and females consistently found these ads to be personally offensive.

With respect to the evaluation of sexual content by sex of the evaluator, the data suggest that men and women vary markedly in their evaluation of both nudity and suggestiveness. These differences become more pronounced as the level of sexual content increases. Considering the evaluations of the nude ads, for instance, it was found that male and female evaluations were almost bi-polar. Males tended to evaluate male nudity poorly while females evaluated male nudity generally positively and vice versa for female nudity. For the male-female nude suggestive ad, females evaluated this portrayal negatively while the male evaluations were generally positive.

To better understand how personal differences affect the evaluation of sexual content, correlation analysis was performed between the evaluations of each ad and the cognitive measures collected in the first experimental session. Correlations were computed for the entire sample and separately for men and women, these latter results being shown in Table 2. Inspection of this data presents some interesting findings. Relative to men, women tended to respond to the task in a much more uniform manner. For women, few correlations were found to be significant at less than or equal to the 0.05 level and no consistent pattern was apparent. However, for men a different picture appears. All three personal difference measures, social values, attitudes toward role portrayals and role orientation, appear to be related to males' evaluations of sexual content. In addition, this relationship becomes more pronounced as sexual explicitness increases. Of these three measures, social values--a measure of the subject's orientation toward personal and social sexual mores--was found to be the most highly correlated with the four affect scales. Men with more liberal social values were more likely to evaluate the "ads" more positively and vice versa. This relationship becomes strongest in the "suggestive" situations and when nudity increased. For example, the twelve correlation coefficients computed between social values and the four affect scales in the three nude portrayals ranged from 0.2191 to 0.3773 and all were highly significant. Although related to a lesser extent than were social values, attitudes toward role portrayals and role orientation were also found to be correlated with the evaluation scales. Men who hold more critical attitudes toward how men and women are portrayed in advertising tended to evaluate the ads more negatively. Role orientation was related in a fashion similar to that of

social values. The more modern or liberated the mens' orientation to social roles, the more positive the evaluation of portrayals presented. Again, these relationships were most pronounced for the more explicit conditions.

TABLE 2

CORRELATIONS BETWEEN PERSONAL DIFFERENCE MEASURES AND ADVERTISING PORTRAYAL EVALUATIONS BY SEX OF RESPONDENT

Por.	Evaluations by Males (n=100)			
	Taste r (Sig.)	Interest r (Sig.)	Appeal r (Sig.)	Offense r (Sig.)
<b>M-FC</b>				
SV <sup>1</sup>	.2860(.002)	.1008	.2316(.010)	.0283
RP <sup>2</sup>	-.0279	-.0607	-.0870	-.1950(.026)
RO <sup>3</sup>	-.1187	-.1418	-.1036	-.0826
<b>F FC</b>				
SV	.1768(.039)	.1259	.0617	.0450
RP	-.0368	.0175	.0005	-.1461
RO	-.2168(.015)	-.0763	-.1556	-.2144(.016)
<b>M-F FC (sug)</b>				
SV	.2175(.015)	.1786(.038)	.1881(.030)	.1471
RP	.2446(.007)	.0850	.0753	.1843(.033)
RO	-.0468	-.0537	-.0580	.0117
<b>M PC</b>				
SV	.2668(.004)	.0527	.2060(.020)	.0664
RP	.1206	.0456	-.0851	.1889(.029)
RO	-.1159	-.0096	-.1771(.039)	-.0486(.315)
<b>F PC</b>				
SV	.0826	.0301	-.0126	.1740(.042)
RP	.1744(.041)	-.1017	-.1018	.0896
RO	.1501	.1668(.049)	.2201(.014)	.0309
<b>M-F PC (sug)</b>				
SV	.2820(.002)	.2662(.004)	.1918(.028)	.1863(.032)
RP	.2141(.016)	-.0187	-.0447	.0623
RO	-.0644	-.0476	-.0147	-.0306
<b>M Nude</b>				
SV	.2820(.002)	.2662(.004)	.1918(.028)	.1863(.032)
RP	.1991(.024)	-.0394	-.0534	-.0354
RO	-.2756(.003)	-.2739(.003)	-.2222(.013)	-.0737
<b>F Nude</b>				
SV	.3473(.001)	.3197(.001)	.3773(.001)	.2607(.004)
RP	.1567	.0193	.0598	.0906
RO	-.1162	-.0545	-.0834	-.0241
<b>M-F Nude</b>				
SV	.3763(.001)	.3722(.001)	.3547(.001)	.2702(.003)
RP	.2267(.012)	.0194	.1927(.027)	.1098
RO	-.2125(.017)	-.1351	-.2131(.017)	-.1248
<b>Evaluations by Females (n=42)</b>				
<b>M FC</b>				
SV	.2121	.1296	.2406	.1426
RP	.0346	-.0956	-.0005	.0272
RO	-.1709	-.1306	-.1441	-.1013
<b>F FC</b>				
SV	.0964	-.0175	-.0338	-.2151
RP	-.0766	.0088	.1752	.0456
RO	-.0338	-.0507	.0973	.0517

<b>M-F FC (sug)</b>				
SV	.0831	-.0757	.1421	-.0564
RP	.1554	-.1006	-.0347	.4412(.002)
RO	.0471	.1918	.0366	.2024
<b>M PC</b>				
SV	.3001(.028)	.3746(.008)	.3066(.024)	.2567(.050)
RP	.0436	.0158	.1123	.2621(.047)
RO	-.2350	-.3015(.028)	-.0992	-.1156
<b>F PC</b>				
SV	.0884	.0885	.0827	-.0481
RP	-.0956	-.1344	-.1222	.0853
RO	-.0453	-.0748	-.0937	-.0368
<b>M-F PC (sug)</b>				
SV	-.2257	.0363	.0306	-.2393
RP	.1049	.0709	.0543	.1671
RO	.0810	.0735	.1376	.2280
<b>M Nude</b>				
SV	-.2257	.0363	.0306	-.2393
RP	-.1834	-.1399	-.2172	.0161
RO	-.2351	-.2123	-.1699	.0439
<b>F Nude</b>				
SV	-.0168	.0811	.0274	-.0576
RP	.0783	.2082	-.0601	.2985(.029)
RO	.0381	.1769	-.0054	.0594
<b>M-F Nude</b>				
SV	.1169	.1870	.0353	.0131
RP	-.0468	.0061	-.0582	.2784(.037)
RO	-.2087	.2766(.038)	-.1047	-.0270

<sup>1</sup>Ewell's Social Value Scale: higher score represents more conservative sexual values.

<sup>2</sup>Attitudes toward role portrayals in advertising: higher score represents more critical attitude toward role portrayals in advertising.

<sup>3</sup>Arnott's Role Orientation Scale: higher score represents more liberated attitude toward social roles.

Conclusions and Implications

This study indicates that men and women vary greatly in their evaluation of sexually related advertising portrayals. In addition, for men these evaluations are strongly related to the individual's personal sexual orientation and to a lesser extent with attitudes toward role portrayals and role orientation. All relationships become strongest when nudity was most explicit and when suggestiveness was present in the portrayals. As sexual content became more explicit, both men and women tended to evaluate portrayals of the opposite sex more highly but tended to evaluate portrayals of the same sex in a generally negative manner.

These results, when compared to those of previous studies, tend to support the conclusion that the use of sex in advertisements must consider situational factors as well as predispositions of the receiver. In fact, while the experiment here differed somewhat from that conducted by Peterson and Kerin, the results indicate that viewers are not always likely to react less favorably toward nudity. Rather, the evaluation of the appropriateness of nude models is likely to be affected by the sex of the model and the perceiver, with each less likely to assign positive evaluations to those depicting the same sex. The necessity of evaluating situational factors is thus obvious. Future research studies should examine factors likely to intervene with

with such evaluations, including message content, media selection, etc.

#### Appendix 1

##### Description of Advertisements

- Advertisement 1: Male Fully Clothed  
Male standing, partially leaning on a fence. Background includes white fence and field. Man's attire is three-piece suit.
- Advertisement 2: Female Fully Clothed  
Female standing with hands on hips. Background consisting of solid blue with no objects or figures. Attire includes modern dress.
- Advertisement 3: Male Partly Nude  
Male apparently leaping for a fence post. Background is of country field variety, thought not clear. Male clothed only in briefs.
- Advertisement 4: Female Partly Nude  
Female standing. Background is vacant. Female clothed only in very sheer lingerie.
- Advertisement 5: Male Fully Nude  
Male standing, arms extended over head in body of water. Frontal view from slightly below waist.
- Advertisement 6: Female Fully Nude  
Female nude combing hair sitting in front of a mirror. Background contains typical bedroom furniture.
- Advertisement 7: Male-Female Fully Clothed (Suggestive)  
Male standing, partially leaning on desk, attired in three-piece suit. Female lying on floor, fully clothed. Dress is pulled to mid-thigh, one leg extended toward male. Background is an office of high quality appearance.
- Advertisement 8: Male-Female Partially Clothed (Suggestive)  
Male and female in body of water, embracing and kissing. Male nude from waste up. Female wearing bathing suit.
- Advertisement 9: Male-Female Nude (Suggestive)  
Male and female fully nude embracing and facing each other. Background red and vacant.

##### References

- Catherine C. Arnott, "Husbands' Attitude and Wives' Commitment to Employment," Journal of Marriage and the Family, November, 1972, pp. 673-687.
- Robert Chestnut, Charles LaChance, and Amy Lubitz, "The 'Decorative' Female Model: Sexual Stimuli and the Recognition of Advertisements," Journal of Advertising, 6 (Fall, 1977), pp. 11-14.
- A. N. Ewell, Jr., "Inventory of Values," in A. Robinson and P. Shaver, eds., Measures of Social Psychological Attitudes, Institute for Social Research, University of Michigan, Ann Arbor, 1972, pp. 457-463.
- Sigmund Freud, as translated by N. Fodor and F. Gaynor in Freudian Dictionary of Psychoanalysis, Greenwich, Conn., 1958.

William J. Lundstrom, and Donald Sciglimpaglia, "Sex Role Portrayals in Marketing," Journal of Marketing, July, 1977, pp. 72-79.

Bruce John Morrison, and Richard C. Sherman, "Who Responds to Sex in Advertising?" Journal of Advertising Research, Vol. 12, April, 1972, pp. 15-19.

"Not Much Sex in Ads, AA Workshoppers Hear," Advertising Age, Vol. 44, August 6, 1973, p. 1ff.

"Nudity is Ad Fad, Not Trend, Four A's Told," Advertising Age, Vol. 39, October 28, 1968, p. 1ff.

Robert Peterson and Roger Kerin, "The Female Role in Advertisements: Some Empirical Evidence," Journal of Marketing, Vol. 41 (October, 1977), pp. 59-63.

Major Steadman, "How Sexy Illustrations Affect Brand Recall," Journal of Advertising Research, Vol. 9, February, 1969, pp. 15-19.

Gordon L. Wise, Alan L. King and J. Paul Merenski, "Reactions to Sexy Ads Vary With Age," Journal of Advertising Research, Vol. 14, August, 1974, pp. 11-16.

HOW SHOULD WOMEN BE PORTRAYED IN ADVERTISEMENTS?--  
A CALL FOR RESEARCH

Mary Lou Roberts, Boston University  
Perri B. Koggan, Grey Advertising, Inc.

Abstract

The roles which women portray in advertisements have been the subject of much public criticism and some empirical research. This paper reviews that research, describes problem areas which need research and presents hypotheses whose exploration should provide useful insights aimed at improving both advertising and over-all marketing strategy.

Introduction

In recent years advertisers and their agencies have received criticism from a wide variety of sources concerning the manner in which women are portrayed in advertisements. The critics state that the women shown in ads are too often "only housewives;" stupid or incompetent; dependent on men; decorative or sex objects; passive; and not involved in making major decisions (Advertising Age, April 21, 1975).

There is no doubt that many of these criticisms were, and to an extent still are, valid. There also is no doubt that many advertisers have attempted to respond by discarding stereotypes and trying to create more appealing role incumbents. If any doubt exists that they have made these attempts in the face of extremely sparse information, we hope to dispell this doubt and will present a list of hypotheses which, though it is by no means exhaustive, is intended to encourage research in this area. Such research should provide guidance for day-to-day advertising decisions and at the same time increase our knowledge of consumer behavior in general and female consumers in particular.

Even though advertisers act in good faith, their efforts at discarding stereotypes while appealing to today's woman may encounter wholly unexpected criticism or marketplace failure. Two examples will illustrate some of the pitfalls.

A cosmetics ad showing the face of an attractive young woman was captioned: "Your face isn't safe in this city." According to a company spokesperson, this was intended to be "a straightforward pitch for skin protection against smog" (Liddick, 1978). This particular advertisement was attacked by Women Against Violence Against Women, a group protesting messages which portray violence directed at women. Both the manufacturer and agency agreed that the possibility of misinterpretation existed and withdrew the ad (Liddick, 1978).

The second example deals more directly with role portrayal. Researchers created four versions of an ad for an instant breakfast drink. The ads varied only in the occupation of the spokesman - a housewife, a grade school teacher, a cab driver and a Ph.D. In an experimental setting the ad with the female Ph.D. was the least preferred. The reason appeared to be that this food product was associated with more traditional female roles (Advertising Age, April 18, 1977).

The problem of female role portrayal in advertisements has many dimensions. We intend to deal with only a selected few of these dimensions -- most viable role for the chief female actor, her relationship to and inter-

action with significant others and the relationship between role portrayal and selected product categories. We have chosen these dimensions because they point to avenues of research which can help marketers in their analysis of basic strategy decisions such as product positioning and market segmentation while they improve their promotional strategy and execution. Some dimensions such as sexual innuendo, nudity, and violence have been omitted from our discussion because, although they are important creative issues, they seem less germane to major strategic considerations.

Before we discuss topics which need to be researched, it is necessary to review the existing empirical work.

Literature Review

Presence of Role Stereotypes in Media Advertising

Courtney and Lockeretz (1971) conducted a content analysis of 729 ads appearing in eight general interest and news magazines during 1970. They found few women shown in employment outside the home. Most employed women were entertainers; none were shown in professional or executive roles. This was true of both ads showing women only and those showing men and women together, although the frequency of employed women increased when men and women were shown together. Women were rarely shown interacting with one another. Their conclusion was that stereotypical portrayals of women were dominant.

A followup study done by Wagner and Banos (1973) used ads from 1972 issues of six of the same magazines (Reader's Digest was omitted and the New York Times Magazine was substituted for Look which had ceased publication). The number of employed women shown had increased from 9 to 21 percent with some women shown as professionals, semi-professionals, sales people and in other white-collar occupations. Fewer women were shown as entertainers or sports figures. However, more non-employed women were shown in a decorative role and fewer were shown in a family or recreational role. There was no change in the frequency of interaction between two or more women, in female involvement in major purchase decisions and in their portrayal in institutional settings. This study, then, presents a mixed picture. Employed women were better represented, although this 21% representation is far from matching the 49% of American women between 15 and 65 employed in 1972 (U.S. Department of Commerce, 1976), but the manner in which non-employed women were portrayed seemed to have degenerated.

Another replication by Belkaoui and Belkaoui (1976) added 268 ads from eight magazines published in January 1958 to the two sets of existing data. This study found that the same stereotyping reported in 1970 and 1972 with regard to employment status, occupational roles and involvement in major purchase decisions also existed in 1958. The percentage of working women shown in ads was slightly higher in 1972 than in 1958, but in all three years studied, the women were largely entertainment and sports personalities and secretarial and clerical workers. The non-working women were even more likely to be portrayed in decorative roles in 1972 and less likely to be shown

in family roles. Their conclusion is that the stereotypes which existed prior to the advent of the Women's Movement were being perpetuated in the mass media of the 1970's.

A more detailed analysis by Sexton and Haberman (1974) included ads published during 1950-51, 1960-61, and 1970-71 covering six product categories -- cigarettes, beverages, automobiles, home appliances, office equipment, and airlines. They evaluated 1,827 ads on eleven dimensions which encompassed the number of persons and types of role portrayals in the ads, their relationships to one another and to the product, and the setting of the ads. Although there was some variation in the findings from one product category to another, the general results agreed with earlier studies in that there was some increase in the number of women shown in working roles from earlier to later years and a decrease in women portrayed in home and family-oriented situations. The frequency with which women were portrayed as rather passive social companions of men was large and stable even though women were shown in more varied situations as time passed. They did not, however, find an increase in decorative portrayals but found that there were more women portrayed as being alluring to men.

Venkatesan and Losco (1975), in a study of magazine ad role portrayal covering the years from 1959 to 1971, again confirmed the existence of sex object, physical beauty and female dependence stereotypes. They, too, noted a decrease in both sex object and family-oriented roles.

Dominick and Rauch (1974) conducted a study of 1,000 prime time television commercials and found that the sex object and housewife/mother role stereotypes also existed in that medium.

#### Effect of Role Stereotypes on Attitudes toward Advertisements

A study by Mazis and Beuttenmuller (1973) using small samples and disguised advertisements found college women reacting to favorable and unfavorable role portrayals in accordance with their positive or negative attitudes toward Women's Liberation.

Wortzel and Frisbie (1974) approached this problem by asking 100 women to choose the role portrayal (neutral, career, family, fashion, sex object) which would most enhance the desirability of ads for seven different product categories. In this sample, preference was expressed for a family-oriented role for both small and large appliances. Where women's grooming and personal care products were concerned, the career role was preferred but the traditional and fashion roles were judged acceptable. Grouping these women by positive or negative attitudes toward Women's Liberation as measured on a 22 item scale did not show any overall differences in preference for role portrayals between the two groups. Instead, the nature of the product category seemed to be the determinant of role preference. They did find, however, that women with positive attitudes toward Women's Liberation were more likely to select a specific role than the neutral portrayal and were more likely to prefer the family-oriented role than were women with negative attitudes. Women who responded positively to a repression sub-scale imbedded in the Liberation scale were more likely to reject any role choice.

A similar study by Duker and Tucker (1977) using a sample of 104 college women and actual print advertisements also found no differences in perception of ads portraying sex-stereotyped roles between women whose orientation to Women's Lib (as measured by the Barron

Test of Independent Judgment) was strong and those whose orientation was weak. They suggest that underlying value structures may be more useful in analyzing women's reaction to role portrayals than are attitudes which are conceptually more superficial and transient.

Most of the available evidence, then, suggests that the presence of sex-role stereotyping in ads may not generate unfavorable perceptions of those ads even in women who hold positive attitudes toward Women's Liberation.

#### Effect of Attitudes Toward Role-Stereotyped Ads on Intention to Buy

Lundstrom and Sciglimpaglia (1977) used a mail questionnaire to study attitudes of 150 women and 114 men toward role portrayal of both men and women in ads, company image of firms whose ads conveyed sexist images, and intention to buy products of such firms. The women in the sample had more negative attitudes toward the realism of the roles portrayed by both women and men. The negative attitudes of the women carried over to company image but no significant difference was found in intention to buy between men and women. Younger, better educated, higher SES and less traditional women held the most critical attitudes toward advertising. However, older women and those from lower income households were most likely to agree that they would discontinue purchasing a product whose ads they found offensive, although they would continue to buy other products made by the same company. Lundstrom and Sciglimpaglia, therefore, conclude that the impact of offensive ads on sales may be slight because the women in their sample who are most critical of advertising also have characteristics which make them potential opinion leaders and/or consumer advocates but have little inclination to translate negative attitudes into avoidance behavior or positive action.

Several conclusions can be drawn from the exhaustive review of empirical studies:

- (1) Sex role stereotypes are present in mass media advertisements and there is no clear evidence that, overall, the frequency of such stereotyping is decreasing.
- (2) Most of the available evidence is based on analyses of print media.
- (3) There is an abundance of data which confirms the existence of stereotypes but little which relates presence of stereotypes to attitudes toward advertisements and even less which deals with the effect of stereotyping on intended purchase behavior.
- (4) The product category itself may be the chief determinant of preferred role portrayal.
- (5) Reaction to role portrayal in advertisements may vary among groups of women with different socio-economic characteristics.
- (6) All published studies to date have used research designs and/or samples which lack total applicability to real world conditions.

#### Where We Stand Now

As we said in the beginning, the advertiser is faced with a multitude of decisions and little published empirical data. Also, some of the data which does exist is contradictory.

A panel appointed by the National Advertising Review Board did examine existing attitudinal studies and

presented checklists of both destructive and constructive portrayals (Advertising Age, April 2, 1975). The intent of this effort was laudatory, but a careful examination reveals that they have essentially recommended that advertisers avoid known stereotypes and present positive female role models in a manner which enhances female self-esteem and encourages the realization of female potential. This is a beginning, but hardly the detailed guidance the advertising strategist needs.

Much research is needed to generate this detailed guidance. A review of a great deal of non-empirically-based literature and observation of ads themselves have identified the major problem areas. With each problem area we present one or more hypotheses. These hypotheses may or may not be confirmed by actual data. Whether confirmed or not, we believe the data will be useful. We hope they will also stimulate research-oriented thinking about problems and/or product categories not discussed.

With each area we also present ads which exemplify the issues. We again caution, however, that the verdict has not yet been made on whether the presence of stereotyped portrayal does adversely affect marketplace behavior.

#### Discussion of Problems and Hypotheses

**Problem:** Ads often depict the woman as a sex object to be admired by a man for his sake, especially in ads for health and beauty products.



The female model in this mouthwash ad is a wholesome young woman. In one picture she is holding the product. In the other picture she is clinging to a rugged young man. Although the model is attractive in her own right, the implication is clearly that she could not have "caught him" without the product's help.

The female model in this next perfume ad has an exotic, sultry aura which is enhanced by the black background. Her lips are seductively pursed and she is whispering to get "someone's" attention.

This ad seems to be implying that the woman uses the scent for the sole purpose of attracting a man. Isn't it possible that women buy perfume to feel better about themselves?



In this ad for a hair coloring product, an attractive young woman, described in the copy as an active working mother, is shown against a pastoral background. The headline and copy clearly state that both the product and the hobby contribute to her sense of well-being.

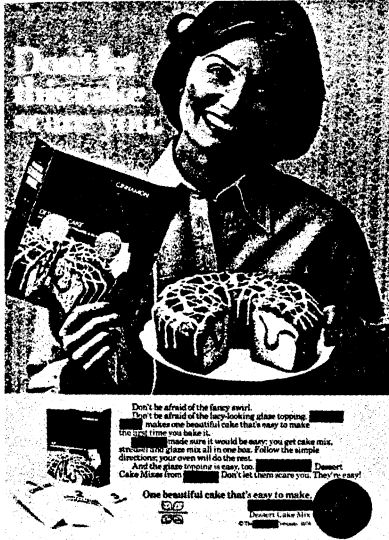


Although this is a very positive role portrayal, the lifestyle may be a bit exaggerated. How many working mothers find time to sit in the middle of a field painting a landscape? Yet it may be a fantasy ad that is very appealing to the target market.

**Hypothesis:** Ads for health and beauty products should appeal to a woman's sense of well being for her own sake -- not to enhance her status as a sex object.

**Problem:** Ads often portray the woman as a helpless fool who needs assistance to help her through the turmoil she has created or seems unable to control.

This next cake mix ad shows another wholesome young woman holding the package and the finished product. Although this ad is not overly offensive, it does portray the woman as being somewhat helpless. Couldn't the product benefit - a cake that is both easy to prepare and attractive - be emphasized without implying that the woman is afraid to try something different?



Don't be afraid of the fancy sort.  
 They're not afraid of the fancy-looking glass topping.  
 makes one beautiful cake that's easy to make.  
 (Just mix the mix with the frosting.)  
 Inside there's a delicious cake mix.  
 And the glass topping is easy, too.  
 Don't let them scare you. They're easy!  
 One beautiful cake that's easy to make.

Actually, this particular problem in role stereotyping is much more prevalent on television. A multitude of examples come to mind -- "ring around the collar," the man in the toilet bowl, masculine "scrubbing bubbles," a "knowledgeable" husband recommending the right furniture polish.



This ad for a washing machine contains both mother and son and emphasizes a special feature of the machine which allows a small load to be washed economically.

This seems to be a positive role portrayal because it shows a sensible young mother performing a necessary household task intelligently.

**Hypotheses:** Whether a woman works or not, she should be shown as competent and creative in using products which help her to perform the tasks her roles and life-style necessitate. These products should be shown helping the woman to achieve objectives that are central in her perception of her role.

: The task-oriented woman (e.g., "I have regular days for washing, cleaning, etc. around the home." "If you want your floors waxed properly, you have to do it by hand."<sup>1</sup>) derives satisfaction from the performance of the tasks themselves.

: The family-oriented woman (e.g., "I put my family's welfare before my own." "I am the guardian of my family's health."<sup>2</sup>) derives satisfaction because she believes she has contributed to her family's well being by the performance of these tasks.

: The woman who is oriented toward activities outside the home (e.g., "I spend a lot of time working for community organizations." "I take pride in my job."<sup>3</sup>) derives her major satisfaction from these activities and, therefore, is attracted by products which simplify in-home tasks.

**Problem:** Ads in which men and women are shown together often use women as decorative or alluring objects.

This ad for glassware shows two attractive young women sharing the celebration of a winning (male) race car driver.

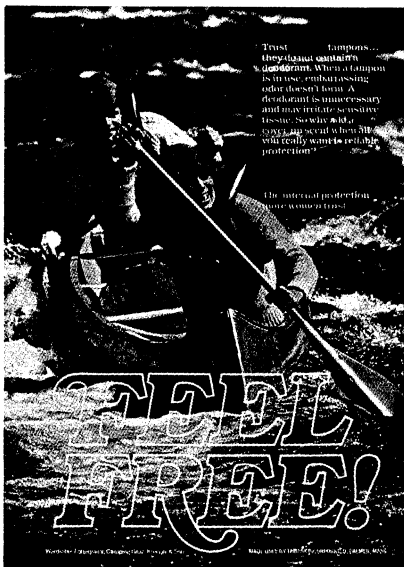
This is an industrial product ad in which all of the models are mere decorations. Obviously, the women shown have not contributed to the man's success, but someone did have to hold the glasses.

This ad for men's cologne shows sexy male and female models in close proximity to one another. While this ad might appeal to a man's fantasy, it was found in a women's magazine. Would this ad compel a woman to purchase the product for "her" man?

<sup>1</sup>Source: Wells, 1971. <sup>2</sup>Ibid. <sup>3</sup>Ibid.



This ad for a refrigerator shows a husband and wife having trouble fitting their old refrigerator into their new kitchen. It is a humorous approach to a real problem. It was chosen as an example because the couple is sharing the dilemma even though the solution is provided by a male celebrity spokesperson.



This ad for a feminine hygiene product shows an attractive young couple canoeing on white water. The contemporary woman wants to pursue an active life whenever she pleases and is probably attracted by themes which convey this life style.

**Hypotheses:** If the target market is men, portraying women as decorative or alluring is appropriate.

: If the target market is traditional women (e.g., "A woman's place is in the home."<sup>4</sup>) or both women and men, portraying the women as equal partners or participants is appropriate.

: If the target market is contemporary women, portraying the women as successful or dominant is appropriate. (e.g., "Women should be free to take jobs

<sup>4</sup> Ibid.

outside the home if they want them."<sup>5</sup>)

**Problem:** Ads in which two or more women appear tend to emphasize competitive interaction.



In this ad for cat litter the hostess of a female bridge group is concerned about cat box odor. Again, the problem is a real one. However, would a group of women -- your friends, at that -- really be this rude?



This ad shows a brand of beer which is bottled in two sizes. A group of women is superimposed on the smaller bottle. These women are enjoying a friendly beer together in a pleasant setting. The quality of the interaction is appropriate to today's woman. The Women's Movement has sensitized her to the pleasures of camaraderie between women.

This ad is also interesting because it deals with a traditionally male-dominated product category but portrays a new version designed to appeal to women. This recognizes the emergence of women as a force in the consumption of these product categories.

**Hypothesis:** Show women as congenial and supportive of one another in a setting appropriate to the life style of the target market.

**Problem:** Advertisers need to identify the type of person who is the most effective spokesperson for their product category.

<sup>5</sup> Ibid.



**"There are three big reasons why washers have stainless steel tubs. No chips. No rust. No snag."**

A stainless steel tub is one way to make your washer more durable. No one else has it. It won't chip or rust. It's not porous. So no rust snags in delicate stuff or knits. And no rust on white clothes.

It backs the stainless steel tub with a limited replacement warranty that lasts for the life of the machine.

On top of that, it also gives the transmission a limited 10-year warranty.

With warranties like these, it's a hard buy to pass up.

**Built better to last longer.**

**Chuck Connors**

Washers are a combination with the best of both worlds. They're built to last. And they're built to be built.

A well-known actor is the spokesperson in this washing machine ad. Why are celebrities chosen as spokespersons for product categories where there is no apparent correlation between their image and the possession of product-related expertise? Even though they may be considered attractive by the target market, are they a credible information source?

**WOMEN IN ENERGY**

**DR. TERRI PECORARO: She's helping more energy out of a barrel of oil.**

To put it simply, after crude oil is refined into gasoline and other products, a tonne residue remains which is usually these byproducts like asphalt.

Dr. Terri Pecoraro, an chemist in the highly technical field of catalysis, is part of a team of scientists working to squeeze more usable energy out of that residue.

As Terri explains it, the right catalysis can help convert the residue into gasoline and heating oil. They can also help remove sulfur and nitrogen to produce cleaner-burning fuels that reduce air pollution - a major goal of research for years.

As a language bonus, catalysis that remove sulfur and nitrogen from petroleum products have been from Ohio, where America's richest storehouse of energy.

There's more to being a chemist than just doing your science, says Terri. Many chemists that look scientifically feasible may not be practical. It's a long down with the chemical engineers and industrial scientists to see if an idea is practical that makes the work intriguing and fun.

Off the job, Terri enjoys bicycling, playing tennis, and swimming along the Jersey Shore with her husband, who is also an chemist.

**Energy for a strong America**

This is an institutional ad with an energy conservation theme. It shows a woman chemist in protective clothing operating highly technical equipment. The woman shown is undoubtedly highly intelligent and expert in her field. Would, however, the average woman be able to identify with this role portrayal?

**Hypotheses:** Traditional women will attribute higher credibility to an authoritative male figure. This is especially true in product categories such as major appliances or those products which represent new technological developments.

: Contemporary women will prefer a female figure who has the necessary level of technical expertise.

: In promoting non-technical products including home and personal care products, all women will

find a female with whose lifestyle they can identify to be the most credible.

**Problem:** Even when the role portrayal attempts to avoid stereotypes, the women shown are often not behaving in a manner appropriate to the role.

**Can top management and word processing get along?**

Not necessarily. Some word processing systems require you to your office with terminals. Reorganize the. Centralize that. Buy one. Buy one. Buy one.

Systems you can get along with. They're flexible. Put them anywhere. Outside your door or in centralized groups. However, your office works. Systems work. Commodity.

And productively enough to please the most cost-conscious management. Types like every word on a. These system for processing. Automatic online type.

at hundreds of words per minute. Magnetic cassettes store your work for later use. With no retyping. And with simple, add-on options. Systems can store information with each other or with computers. And work alone for hours. And send work to other computers. And process text back. And more.

Find out how well you'll get along. Call us now for a demonstration, or for more information. 800-548-3550.

**World Processing System**

This ad for office systems shows a greying female executive seated at her desk. She is looking at a cathode ray tube while being served coffee by a man. Showing a female executive is commendable. However, her facial expression is clearly not that of an executive. In addition, the man serving coffee has no relation to the purpose of the ad, unless it is to exemplify the role reversal.

**"I CARRY CLOUT"**

Clout gives you acceptance.

Clout makes playing Santa easy. Clout opens doors. Clout is a card. Your personal account for Christmas shopping at over 2 million places. All around the world. That's clout. That's Relax.

This ad for a charge card shows a woman shopper holding the card. Since women have traditionally been denied equal access to financial services, the ad's basic message -- the power conferred by the charge card -- is appealing. However, the woman's expression is not ter-

ribly attractive. In fact, it seems to convey the stereotype of a compulsive female purchaser.

It seems desirable to portray women in a variety of roles which would represent the diversity of modern lifestyles. Care should be taken to match role and incumbent and to create situations with which the target audience can identify.

Hypothesis: The intended role portrayal is most effective if all elements of the ad are in harmony with the role portrayal.

#### Conclusion

The question of how to most effectively portray women in all advertising media will not be answered easily or quickly. It does, however, offer tremendous scope for meaningful research and positive action by advertisers and agencies. We hope the issues raised here will be useful in generating both.

#### References

"Advertising Portraying or Directed to Women." Advertising Age (April 21, 1975), pp. 72, 75, 78.

Ahmed Belkaoui and Janice Belkaoui. "A Comparative Analysis of the Roles Portrayed by Women in Print Advertisements: 1958, 1970, 1972," Journal of Marketing Research, Vol. 13 (May, 1976), pp. 168-172.

Alice Courtney and Sarah Wernick Lockeretz. "A Woman's Place: An Analysis of Roles Portrayed by Women in Magazine Advertisements," Journal of Marketing Research, Vol. 8 (February, 1971), pp. 92-95.

Joseph P. Dominick and Gail E. Rauch. "The Image of Women in Network TV Commercials," Journal of Broadcasting, (Summer, 1974), pp. 259-265.

Jacob M. Ducker and Lewis R. Tucker, Jr. "Women's Libbers Versus Independent Women: A Study of Preferences for Women's Roles in Advertisements," Journal of Marketing Research, Vol. 14 (November, 1977), pp. 469-475.

Betty Liddick. "Feminists Go To Battle Against Ads That Victimize Women." Boston Globe, February 22, 1978, p. 1.

William J. Lundstrom and Donald Scigliompaglia. "Sex Role Portrayals in Advertising," Journal of Marketing, Vol. 41 (July, 1977), pp. 72-79.

Michael B. Mazis and Marilyn Beuttenmuller. "Attitudes Toward Women's Liberation and Perception of Advertisements," Advances in Consumer Research, Vol. 1, 1973, 428-434.

Donald E. Sexton and Phyllis Haberman. "Women in Magazine Advertisements," Journal of Advertising Research, Vol. 13 (August, 1974), pp. 41-46.

"'Traditional' TV Women Bother Men, Study Tells," Advertising Age, April 18, 1977.

Louis Wagner and Janic B. Banos. "A Woman's Place: A Follow-Up Analysis of the Roles Portrayed by Women in Magazine Advertisements," Journal of Marketing Research, Vol. 10 (May, 1973), pp. 213-214.

William D. Wells. "AIO Item Library," Chicago, 1971. (Mimeographed)

Lawrence H. Wortzel and John M. Frisbie. "Woman's Role Portrayal Preferences in Advertisements: An Empirical Study." Journal of Marketing, Vol. 28 (October, 1974), pp. 41-46.

M. Venkatesan and Jean Losco, "Women in Magazine Ads: 1959-71," Journal of Advertising Research, Vol. 15 (October, 1975), pp. 49-54.

U.S. Department of Commerce, Bureau of the Census, A Statistical Portrait of Women in the U.S. Washington, D.C.: Government Printing Office, 1976.

SEX ROLES, SEX, AND STEREOTYPING IN ADVERTISING:  
MORE QUESTIONS THAN ANSWERS

Mark I. Alpert, The University of Texas at Austin

Introduction

Recently there has been an increase in attention focused on sex role portrayal and sexuality in advertising. As this session on "Sex in Advertising" illustrates, this attention has come not only from critics of advertising, but from practitioners, regulators, and academicians, all with a variety of perspectives. The papers in this session range from a very specific study of aspects of sex in advertising, to a general theoretical discussion and literature review of sex in advertising, to a more general discussion of issues in role portrayal and stereotyping of women in advertising. Their orientations vary from pragmatic and managerial issues of what attitudinal effects result from sex and nudity in ads, to questions of the empirical validity of various normative statements concerning how advertisers "should" deal with stereotypes in a changing social milieu. Taken as a whole, these three papers provide a useful review of what is thought to be true, suggest issues where questions remain substantial, and propose a number of approaches to studying the phenomena in question.

Let us proceed to examine some contributions from each of these papers, as well as questions that might be raised in regard to each. In closing, we shall discuss issues in the general area of sex, sex roles, and stereotyping in advertising.

First Paper

The first paper, "The Role of Sexually-Oriented Stimuli in Advertising: Theory and Literature Review," by Wilson and Moore, (1979), provides a useful statement of theoretical considerations underlying the differential effects of degree of advertisement sexuality on varying products and audiences. This paper is particularly useful, given the validity of the authors' assertion that most marketing research into this topic has lacked theoretical bases for generating and testing hypotheses. Their theoretical discussion is brief, and space in these papers is limited. However, it illuminates several important elements in the process by which sexually-oriented ads may interact with particular products and receivers to produce particular effects such as attention, retention, recall of brand vs. ad, and attitudes towards the ad, product, and company.

The discussion of arousal is especially interesting, as the authors argue that erotic ads may create disequilibrium, then psychological tension which may cause increased cognitive activity directed towards the ad and/or the advertisement product. Depending on how this advanced level of information processing interacts with the consumer's value system, positive or negative effects may be obtained.

The authors also claim that arousal theory explains the negative impact of sexy advertising, where the inclusion of a female model is not logically related to the product. They reason that since arousal inhibits the impact of irrelevant information, gratuitous use of sexy models may depress perception of advertising messages. However, the logic here is not clear, since it would seem that the irrelevant information that would be depressed would pertain to the sexy stimulus, but not to the advertising messages, assuming they are comparable to those used in the non-sexy ads.

Since descriptive studies rarely have controlled for message content, comparisons of attitudes towards sexy vs. non-sexy hardware or computer ads may have confounded message with degree of erotic content. The authors' explanations for this phenomenon that are based on self-concept theory seem more valid. Here they argue that females who identify positively with a sexy model will like sexy ads, while those who do not want to be seen as sex objects will react negatively to such ads, particularly when the irrelevance of the model to the product enhances perceptions that she is solely used as a sex object.

Discussion of selective perception and absolute limits of sensory stimulation is also effectively illustrated with examples relating how older women may not respond to sexual innuendo in some advertising, due to non-familiarity with the verbal idioms (and lower sensitivity to sexuality). It would be useful to consider the concept of perceptual threshold in the light of subliminal sexuality in advertising. It might be argued that erotic stimuli may not be consciously perceived as sexy by segments of the audience who may nonetheless be positively (or in some cases, negatively) influenced by the subconsciously perceived sexual content. Certainly many advertisers and critics of advertising operate on this premise. Research questions might be formulated in which one could contrast the perceived level of erotic content with that measured by physiological means (e.g. pupil dilation, heart-rate, etc.) as stimulus intensity is varied, and the resulting impact on attitudes and behaviors towards the ads and products, and their covariation with audience variables such as education, values, sex, age, and marital status.

It is worth noting that the theoretical discussion is written mainly from the standpoint of males' perceptions of sexy women in advertisements, with little emphasis on women's perceptions of sexy males. The generality of the phenomenon should be kept in mind, however, and the authors' discussion of empirical findings in terms of both sexes' responses to stimuli of the opposite and/or same sex show that they are aware of this generality.

The final theoretical link is to aggression theory, and the authors direct most of their attention to whether sexy ads provoke aggressive behavior. From a social and public policy, this is a key issue and it warrants concern, but the managerial significance of the sex-aggression link seems slighted. A number of advertisements seem to combine an appeal to sex and aggression drives linked to product characteristics: witness advertising for "Cougar" automobiles featuring Farrah Fawcett-Majors. If research were to establish that those who are most apt to perceive (via selective perception) these ads were also seeking a manifestation for aggressive motivations (or at least tend towards aggressiveness) managers could see relevance to matching erotic ads to "aggressive" products and customers.

The literature review relates the theoretical rationale(s) to empirical findings from psychology and marketing. Considering the vast body of psychological literature to be summarized, the authors did a creditable job. It would have been more useful had they been able to provide the same degree of critical evaluation to this literature that they did for the marketing studies. For example, it would be easier to generalize from

particular studies of differential effects of erotic content on males vs. females if the psychological literature review were to note which, if any, used adult samples. The possibility exists that part of the reason why Kinsey found females "less responsive" to erotic content than did some later studies may be that he used adults, and several later studies used college students, whose females may be more "liberated" and less sexually inhibited than the adult females to whom their responses are perhaps being generalized. It is difficult to comment in depth on a vast literature, but an attempt to account for the effect of sample biases would aid interpretation of generalizations that might be made.

In a related vein the authors should be credited with making the useful distinction between the degree of explicitness in sexual stimuli in psychological experiments (high) vs. that for advertising (low to moderate). The extent to which these differences temper the applicability of psychologically-based studies on advertising predictions is effectively noted.

Six published studies of sexual stimuli in advertisements are reviewed in detail, and an effective table summarizes the variables studied, context, sample, and other relevant considerations. In general, the authors seem on sure ground in this review, and they comment effectively concerning what each study shows and what it does not. Space constrains me from summarizing key findings that are in the paper; rather, I would like to point out areas for further consideration, with the understanding that these comments should not imply that considerable insight has not already been shown by the authors.

In reviewing Alexander and Judd (1978), they point out that nudity and even the "human element" in an ad may distract subjects from learning brand names (confirming distraction hypotheses); yet in a realistic setting they state that net learning may be enhanced due to higher attention to the ads. This is a useful assertion regarding the external validity limitations of showing people ads vs. measuring actual attention, learning, and so forth, in a natural setting (ala Scherwin or K-Tel). One should note that in this study, respondents were only allowed to look for 8 seconds or 15 seconds, with this controlled on a slide projector. It would also be useful to note the necessity of exploring the generality of the nudity - recall relationships over other products and subjects, since Alexander and Judd used college students.

The review of Peterson and Kerin (1977) notes the effective use of experimental design to explore relationships among product type, degree of model nudity, and respondent sex on perceptions of product, ad, and company. Overlooked in their critique of this study (and perhaps relevant to others) is that attitudes towards the "topless" model might have been lower than the less-nude poses because of an interaction with the attractiveness (or lack of) her unclothed figure. Considerable literature suggests that attractiveness affects communication, liking, and so forth; hence erotic stimuli studies may be confounded by the attractiveness of available models. Ratings of this variable might be obtained and its impact on key variables could be measured.

The authors close with some research recommendations, including the use of focus groups for hypothesis generation, multiple measures of advertising effectiveness, and controls for realistic advertisements, without confounding lighting and other factors. Additional research recommendations would make the paper a bit more useful, particularly regarding the hypotheses that should be tested to validate the various theoretical positions that the authors summarized. Suggestions for operationalizing and testing appear in the paper, and their quality is such that even more would be welcome. The paper has

synthesized considerable literature, and yet one has the sense that there is not much that can be generalized at this point. To the extent that more of the theoretical links that are initiated in their paper begin to influence future research, future summaries might become more definitive and more satisfying.

#### Second Paper

The second paper, "Demographic and Cognitive Factors Influencing Viewers Evaluations of 'Sexy' Advertisements" by Sciglimpaglia, Belch, and Cain, (1979), provides empirical evidence concerning some of the conceptual questions raised in the first paper. The authors studied the impact of variations in nudity, "suggestiveness", and sex of the model(s) on perceptions of simulated ads given by male and female students, with varying sexual values and role perceptions. Their findings provide some insight into the relationships among respondent variables and perceptions of "sexy" advertising.

The study is an ambitious one in several respects. The authors randomly ordered 8 dummy ads (with sunsets, children, etc.) and 9 treatment ads that were largely drawn from actual print media advertising examples of varying degree of nudity and suggestiveness. Attitudes towards these advertising excerpts might have more external validity than the usual studies employing "made-up" advertisements. Unfortunately, this realism had to be achieved at the cost of confounding degree of nudity and suggestiveness with other factors that varied nonsystematically across the treatments. Different models were used in each ad, while ideally the same male and female would have been used in a controlled experiment. There are also differences in background that may in some cases have influenced the findings. For example, how much of the female respondents' greater antipathy towards the fully nude female "ad" than the male nude "ad" might be due to the former model being shown in a bedroom setting and the male standing in water and viewed from slightly below the waist? Attractiveness of the models, facial expression, and other confounding variables present additional problems, especially with variations in the models. In that their conclusions generally support those of other studies, this confounding may not be a major problem, but there remain questions of what was causing the effects that were measured.

Another useful extension in this paper is the use of measures of sexual attitude, advertising role portrayals, and role orientations to predict attitudes towards particular ad treatments. For the males in the study, a number of intuitively reasonable and statistically significant correlations were found, particularly those relating conservatism towards sexual values with disapproval of nudity or suggestiveness in ads involving one or both sexes. Incidentally, the correlations in Table 2 would be easier to interpret if scales were scored so that the same sign of a correlation coefficient would have the same meaning across scales. For example, role orientation scores are reversed, such that a negative correlation means that conservative attitude toward social roles is associated with negative statements about the ads, while the opposite is meant for negative correlations with the sexual value scale.

Apparently, these attitude scales were less successful in "explaining" female advertising perceptions, for as the authors noted, few significant correlations were shown. Lest anyone attempt to generalize from the female half of Table 2, it should be noted that the 10 correlations significant at the .05-level of alpha may not be significantly more than the 5 (5% of 108 "tests") that would be expected due to type-I errors, and the variance would be higher due to inter-correlations among variables and ratings of multiple ads per subject.

Response effects may also be present in this study, given that subjects were asked to provide paper-and-pencil statements about advertising that was clearly related to sex and nudity. This is a problem in any subjective statements about advertising, and attitudes towards the products would be better, particularly if obtained less obtrusively. In this case the purpose may have been particularly obvious, given the presence of 7 out of 17 ads that were sexual, and 3 with outright nudity. In the Peterson/Kerin (1977) study, only one nudity treatment per subject was used, possibly making the research objective less obvious, particularly for the less blatant ads (for the "topless" treatment, they too may have this problem). In the present study, response set bias may be amplified by administering the (sensitizing) sexual attitude questions prior to the ads. According to Bem (1972) and Reingen (1978) self-perception theory may predict that having described oneself as sexually liberal, one may respond more positively to evaluation of sexy ads, and negatively for self-described conservatives. Less reactive dependent variables and/or separate measurement of sexual and role attitudes would improve the study.

In summary, the study suggests that prior research showing females less tolerant of sexuality in advertising may be given additional support by their findings that extend the literature to look at suggestiveness as well as nudity. Caution is needed in generalizing the results, due in part to the use of Southern California college students, and also to confounds and other threats to internal validity. The scales that were used may be studied further and their reliabilities and validities strengthened. Further, one might temper statements to the effect that evaluations are "strongly" related to men's personal sexual orientation, since statistical significance may differ from practical significance. Even for nude portrayals, the correlations that were "highly significant" would have  $r^2$ s of .048 to .142 (without adjusting for response-effect's upward bias). From a conceptual standpoint, this study might be integrated with the theoretical positions taken in the Wilson/Moore paper, and it would support some of the conclusions from the literature. No doubt, injecting realism into advertising research in this area can be worthwhile, and future research can build upon the exploratory work here presented. Replications under conditions in which non-sexual aspects of the stimuli are more controlled may add validation to what is here, and careful manipulation of other advertising elements may also be undertaken.

### Third Paper

The third paper, "How Should Women Be Portrayed In Advertisements?--A Call For Research," by Roberts and Koggan (1979) applies a basically case-study approach to current advertising in order to generate some hypotheses about women's role portrayals and their impact on advertising messages. It would have been helpful, in the context of this session, had the authors chosen to deal more extensively with the issue of nudity in advertising, so that one could contrast their views with those of the two preceding papers. From their discussion of the first "problem," one may infer that advertising using women as sex-objects should be replaced by those appealing "to a woman's sense of well being for her own sake." (Roberts and Koggan, 1979). However, here the authors are directing their messages to advertisers aiming at women customers, and it would be interesting to get their views on the ethics and pragmatics of advertisements aimed at males.

Given their scope, the authors provide some useful evidence and arguments on the topic of stereotyping females in advertising. Their treatment of female role portrayal relationships to significant others as depicted in

ads, and relationships to various product categories reflects concern with broader and more fundamental issues than the question of sexy advertising (which is a special case of role portrayal). Their literature review helps summarize the studies of the extent to which women's stereotyping in advertising is increasing or decreasing. Unfortunately, the results of the studies are mixed, and it is not clear whether the "gain" in portrayal in professional roles is offset by an increase in sex-object roles. Studies of attitudes towards advertisements were summarized and found to relate to respondents' attitudes towards women, and socio-economic characteristics. Limitations in the external validity of these studies were noted in a general sense. The literature is well summarized by the authors, who point out that many proposed suggestions for changing the traditional female stereotypes do not provide substantial guidance for advertising strategy. They noted in their introduction that well-meaning attempts to portray women in a positive role may have stray connotations that disturb some women's groups. Others may attempt to place women in a highly professional light but miss the target market of housewives, as when a Ph.D. female biochemist endorses a breakfast drink.

Accordingly, their call for research to test propositions concerning the impact of stereotyped female role portrayal on marketplace behavior is well taken. The balance of their paper is given to discussion of a series of "problems" in stereotyping, examples of advertisements that may be part of the "problem" and/or part of the "solution," and hypotheses concerning how role portrayals should be done. Many of these are interesting examples and speak to alleged problems in the area. It would be useful to know how representative are the ads that have been abstracted for the paper, although a detailed and random sampling of advertisements is probably beyond the scope of the paper.

A more substantial limitation to the usefulness of the paper is the lack of a tight theoretical basis for describing how stereotyping works, why advertisers use it, and how attitudes towards stereotyped females may be likely to vary according to particular sociodemographic, cultural, and other relevant attributes of the (male and female) perceivers of advertising messages. A number of hypotheses are drawn from an implicit framework that would be more effective if made more explicit.

Another area where the paper could become more useful would be through the addition of empirical guides for operationalizing and testing at least some of the specific hypotheses that are generated from the advertising examples presented. For example, if the authors believe that ads for health and beauty products should appeal to a woman's sense of well being rather than her status as a sex object, it would be useful to present a conceptual argument that says for what kinds of women this might be true, and where it might not be. In addition, they might propose a study in which advertisements could be scored according to objective criteria and rated for "sense of well being", "sex object", and so forth, and exposed to subjects who are measured according to their buyer characteristics. Measuring instruments for key advertising effectiveness variables should be discussed, both verbal and more direct measures of attention, interest, and persuasion. Analytical procedures such as analysis-at-variance and covariance could be presented in the context of a proposed study. Some of these measures are treated in the Wilson/Moore paper, and a classic study by Evans (1957) illustrates how ads could be scored for congruence with strategies suggested by motivation research and this measure correlated with readership scores. As far as the present paper is concerned, some of the hypotheses could be tested with Starch scores for males and females.

Although ideally all elements of the comparison ads would be the same except for the role portrayal utilized, this is hard to achieve in an ex post design with current advertisements. However, this paper's exploratory research status might support such analyses for one or two hypotheses, and it would be stronger than the pure case study approach used alone.

Most of the examples and hypotheses consist of providing a rationale for why an ad will be effective, or will not be, in terms of how stereotyping is handled in the ad. The logic is often attractive, but the contribution would be more substantial if we had a stronger conceptual framework, and either more empirical support for the hypotheses, or a proposed methodology for operationalizing and testing the hypotheses that are put forth. A number of the authors' hypotheses are reasonable and interesting; others may be more viewed as controversial, depending on one's perspective. It is hoped that the authors and others may be motivated to continue the work presented in this paper, and thereby provide the empirical validation that is needed to improve our understanding of this area.

#### Discussion

The papers in this session are certainly of topical interest. Beyond the matter of sexy advertising may be more fundamental questions of managerial strategy and responsibility. How can advertisers creatively respond to the challenges of forsaking traditional sex-role stereotypes, and what impact will this have on advertising to customers whose sex and sex-role orientations may vary? Much work remains to be done to provide theoretical and empirical understanding of the extent to which advertising uses stereotypes vs. the extent to which it creates or at least reinforces them. Most of the papers have focused on female role-typing in advertising, but male role-typing is at least implied. The societal and managerial perspectives on what advertisers should do in response to changes in sex role orientations and sensitivities to them will continue to be debated for some time to come. Further, if Tucker (1976) is correct in the assertion that verbal attitudes towards these issues may be discrepant from behavior, research in this area will be more challenging and potentially more misleading than in many other problem areas.

A number of implications for research are covered in or may be stimulated by the papers discussed in this session. Wilson and Moore have presented a first step in building a conceptual framework for the analysis of sexually-oriented advertising. The broader sex-role issues may require integration of recent contributions from sociology as well.

A number of research issues have been noted above by the session authors and in these comments. Among them is the choice of dependent variables to measure advertising effectiveness. Most of the marketing literature has used statements towards the advertisements. Less reactive measures that utilize other indices of interest, attention, attitude change, and behavioral intentions should be employed in future research and validity-generalizations may be attempted from these multiple measures. Further, respondent sensitivity to this topic calls for care in securing non-biased samples, asking questions (Blair, Sudman, Bradburn, and Stocking, 1977). and in debriefing.

The stimuli variables to be operationalized and explored also need research under controlled conditions. Among these are the number and sex of persons used as models in the ad, advertising medium, degree of nudity, degree of attractiveness of the model(s) and/or their clothing, matching of sex of model with that of ad respondent (or non-matching), background, product or service advertised

brand "name", perceived relevance of model to advertising situation, and numerous interactions among these related stimulus variables. For example, consider the effect of partial nudity on the meaning of brand names such as "Close-Up" vs. "Crest" toothpaste.

Similar questions remain concerning the differential effects of stimulus variables on receivers of varying sex, role-orientation, sexual attitudes, involvement with the product, and a variety of psychographic variables. Situational variables also may affect the communication process, with variations induced by who else is present during the advertising exposure, individual vs. group decision-making, and so forth. In addition, there are a number of ethical issues in conducting such research, as well as questions concerning the legitimate use of such findings in a commercial context. No matter what perspective from which one approaches the area, improved knowledge of the process by which sex roles are communicated and to what effect is a prerequisite to effective actions to influence advertising strategies. At this point, most of what is believed is fragmented and anecdotal. If future research is done with the care suggested by a careful study of the existing literature, our knowledge may be considerably strengthened by those efforts

#### References

- M. Wayne Alexander and Ben Judd, Jr., "Do Nudes in Advertisements Enhance Brand Recall?", Journal of Advertising Research, 18(February 1978), 47-50.
- Donald J. Bem, "Self-Perception Theory," in L. Berkowitz (ed.), Advances in Experimental Social Psychology, Vol. 6 (New York: Academic Press, 1972)
- Ed Blair, Seymour Sudman, Norman M. Bradburn, and Carol Stocking, "How to Ask Questions About Drinking and Sex: Response Effects in Measuring Consumer Behavior," Journal of Marketing Research, 14(August 1977), pp. 316-321.
- Franklin B. Evans, "Motivation Research and Advertising Readership," Journal of Business, 30(April 1957), 141-146.
- Robert Peterson and Roger Kerin, "The Female Role in Advertisements: Some Empirical Evidence," Journal of Marketing, 41(October 1977), pp. 59-63.
- Peter H. Reingen, "On Inducing Compliance With Requests," Journal of Consumer Research, 5(September 1978), 96-102.
- Mary Lou Roberts and Perri B. Koggan, "How Should Women Be Portrayed in Advertisements?--A Call for Research," in Advances in Consumer Research, 6(Chicago: Association for Consumer Research, 1979).
- Horace S. Schwerin, "Why Television Commercials Succeed," in Robert Ferber and Hugh G. Wales (eds.), Motivation and Market Behavior, (Homewood, Illinois: Richard D. Irwin, Inc., 1958), pp. 321-333.
- Donald Sciglimpaglia, Michael A. Belch, and Richard F. Cain, Jr., "Demographic and Cognitive Factors Influencing Viewers Evaluations of 'Sexy' Advertisements," in Advances in Consumer Research, 6(Chicago: Association for Consumer Research, 1979).
- W.T. Tucker, "A Long Day of Discrepant Behavior," in Marketing: 1776-1976 and Beyond, edited by Kenneth L. Bernhardt, (Chicago: American Marketing Association, 1976), 351-353.

R. Dale Wilson and Noreen K. Moore, "The Role of Sexually-Oriented Stimuli in Advertising: Theory and Literature Review," in Advances in Consumer Research, 6 (Chicago: Association for Consumer Research, 1979).

## PRISONER OF SEX IN ADVERTISING

Edward C. Strong, Tulane University

At the risk of offending my literate friends in the Association and alienating everyone else now gratuitously identified as illiterate, I will point out the heritage of the title of this piece. In 1971, Norman Mailer felt himself sufficiently moved by the activities of the Ms.es Millet, Friedan, et. al., to make a clear exposition of his position in regard to men, women, sex and liberation. This is to say that his exposition was clear to me, although it may have fallen far short of clarifying things for the sisterhood. I contemplate a similar fate for my own efforts with an equanimity approaching that of Mailer. In any event, those desirous of tracing the lineal descendency of this article may find its roots in Prisoner of Sex--and judge the soundness of its parentage thereby.

This article should serve as a hot, wet kiss in the ear of ACR, delivered with all the sincerity and feeling which the recipient will allow me to offer, a perhaps vain attempt to rekindle some old understanding that we all have had in our fumbling dealings with the subject of sex. Perhaps, if this session were entitled "Humor in Advertising," we could expect to laugh a little while perusing the papers or during the discussion. Why, then, in this session do we apparently have no expectation of arousal or even the more tender emotions that allay themselves with sex? The papers are cold and rational things, with only one tiny wink of a leer, and a voyeuristic one at that, showing through. This comes in the next-to-last paragraph preceding the conclusions of the Wilson and Moore piece. Following a thorough, but academic, review of theoretical notions which may be useful in the academic study of sex in advertising and a cataloguing of empirical findings on the subject, we are suddenly faced with a reference to "a measurement device for female sexual arousal called photoplethysmographic probe." What a name! What intricacies of stainless steel, wiring, and pneumatic adjusters it conjures up? Just to own one would remove a researcher's work from the ordinary and deliver it to a more convoluted plane. And the applications! What experiments couldn't be designed with such a device to keep track of the subconscious reactions?

Not that Wilson and Moore are being lauded for this momentary lapse into titillation; I am merely singling out what is the only place, for me, in these papers about the subject of sex in which the subject is momentarily suffused with meaning. I would go so far as to conjecture that Wilson and Moore have incurred the wrath of Roberts and Koggan with their reference to the instrumentation of the erotic laboratory. What rejoicing, after all, should accompany the metrification of the arousal of the female if this signifies another foray by the male to subject those of the feminine persuasion to the indignity of unequal treatment. Even the unliberated male is left to wonder whether or not the male equivalent of the photoplethysmographic probe lies throbbing in the dark recesses of an inventor's mind, or perhaps, in autoclaved pristine glory, it awaits the magic of a dense polysyllabic name before its introduction to an adoring public.

The lack of sex in academic treatments of sex may strike the reader as entirely appropriate, and, the author added sardonically, in keeping with some of our more cherished academic traditions. And while my stated goal is to bring about arousal, it is not my desire to stir the passions with pages of steamy prose. The point is

that from my point of view, all of the papers for this session seem to lack a sensitivity about the size and scope of their concern. Sex is a huge and mythic issue, not one to be dealt with through experimentation on 142 university students, nor through a review of theoretical constructs that may assist in explaining its effectiveness in advertising, nor by reviewing the role models portrayed in a handful of advertisements. Sex in advertising is a pervasive thing, a large lolling maja of an issue, and nibbling at the earlobe or sucking pensively at a toe is merely toying with the lady. Let us be serious in our intentions, be they good or evil.

Sex in advertising cannot be contained by some simple definition, circumscribed to some set of variables concerning, for example, the graphic explicitness of illustrations. One would have thought that Freud had been sufficient to make everyone aware of the extent to which sex permeates our characters and our actions. And if Freud had been sufficient for mankind in general, then Dichter would have sufficed in marketing to alert us to our more unconscious consumptions. His sexual interpretation of the problems with Saran Wrap comes to mind. In her housewifely role, a homemaker's concern with the well-being of her family extends through feeding them to the management of the process of cooking and raw materials procurement and maintenance. According to motivational research, the homemaker contemplating the contents of the refrigerator is disturbed by the nakedness of foodstuffs covered with clear plastic wrap. This nakedness is experienced on several levels; there are the individual's own feelings about nakedness (do nudists use Saran Wrap more willingly than persons less used to dealing with things in the buff?); there are the concerns about lack of protection for the items themselves; and, through the process by which they are to serve the needs of the family, the lack of protection for the family itself. Sex in the Refrigerator! The Secret Lives of Peeled Vegetables and Yesterday's Leftovers! Yes, this is ludicrous, yet at some level it is a dead serious indication of the extent to which sex, and in particular our own individual manner of coping with sex, is one of the few continuing undercurrents of our existences. Let us recognize, then, that when we speak of sex in advertising we speak of all advertising and all advertisements; not just Joe Namath or Pete Rose being paid enough to make a laughingstock of masculinity, but the less obvious ones as well.

### The Gender of Advertising

Users of the English language are, I have felt for some time, unfortunately vague about gender. While other languages helpfully provide a sex assignment for objects, about the only one that we seem to have acquired from the English (in their mania for the sea) is that a ship is feminine. Hardly enough to develop a keen sense of gender. I would contend that the gender of advertising is primarily female. By this I do not mean that an advertisement is simply a feminine object, but rather that the intent and behavior of an advertisement is basically feminine. Advertising is not the kind of a tool which one can use with immediate effect; its objectives are, by and large, for the longer term. To achieve their effects, advertisements comport themselves to insinuate themselves into the minds of the viewer, they cajole, they remind, they seduce, and, at times, they even nag.



I have no doubt that I now face two sorts of hostility; one from readers who find it difficult to grasp what on earth I mean by the gender of advertising, and another from readers who feel that I have done the ladies an injustice by my characterization of their behaviours, another chauvinist beating his drum. To the first group, I offer this further explanation: In the classic economist's view, the role of advertising is to provide factual information about a product in order to enable the consumer to make rational decisions in achieving satisfaction of his need to equate marginal utilities. This view of advertising envisions a masculine advertisement: A blunt statement of fact, a take-it-or-leave-it, no-frills, unsubtle approach. As far as I can tell, this type of advertising has only existed in the minds of economists.

To the other group, I can only offer cold comfort. Why the Creator stopped at only two sexes, I will never know; Admiral Byrd seemed to show much better judgment in his plans for surviving at the South Pole in his insistence that there be either three people in the party or one person alone. Nevertheless, here we are, with only two sexes and their permutations, and I am told by a female biologist of some standing, there are basic physiological and psychological differences between the two which make my statements about the comportment of the female acceptable in a general way.

The notion of the gender of advertising is useful in the consideration of Sex in Advertising. Even after removing innuendo and any graphic representation, we are left with the notion that basically there is a gender, a comportment of advertising which can be characterized as based upon sex. It cannot be legislated away nor can it, in all likelihood, be empirically observed. And in pursuit of its objectives it will put as many strings to its bow as are required, including the use of sex itself; Diana the Huntress portrayed with one breast bared.

#### Issues

Our concerns with Sex in Advertising may stem from several interests, only some of which make any sense from the point of view of a marketer. I will review some of these under the headings of Advertisements as Erotica; Controlling Sex in Advertising; and How (Sex in) Advertising Works.

#### Advertisements as Erotica

The Wilson and Moore article, as noted above, makes reference to the borrowing of equipment from the researchers on erotica to conduct research on advertising, and in a less explicit way the Scigliompaglia, Belch, and Cain article's selection of stimuli tends in a similar direction with the inclusion of a "suggestive" stimulus. The latter article, of course, attempts to document how variation in the stimulus affects the persuasive ability of the advertisement. From what I have seen of advertising, the level of explicit sexual stimulation afforded us in our daily reading and viewing is abysmally low. If we depended upon advertising to arouse us, the race would soon die out.

If the researcher wants to stick it out with erotica, he seems well advised not to try to find it in advertising, and I think we can save ourselves a lot of grief and studies with ambiguous outcomes if we leave that approach alone. On the other hand, attending to my contention about the gender of advertising, there would seem to be a lot more implicit sexual stimulation that goes on in advertising than explicit stimulation. But this kind of stimulation does not bring on heavy breathing and is going to require much more than imaginative

gadgetry in its documentation. Therefore, it seems to me quite justifiable on our parts to turn our backs on the erotic laboratory in our further efforts.

#### Controlling Sex in Advertising

The other day in the hallway I came across a student showing a professor a cartoon related to a paper she was doing for the Business Policy class. Knowing that the chief executive of the company depicted in the case would be visiting and might read some of the papers, I suggested that she not attach the cartoon to her paper, as she apparently was going to do since the cartoon had been carefully mounted on a sheet of paper. She replied that she definitely would not and that it had only been mounted on paper to cover up the back of the cartoon, which had come from *Playboy*. Interested in this old-fashioned display of a sense of modesty, I asked her whether she really felt that she had to protect her fellow graduate students to that extent. To which she quickly replied, "Oh, no, but we do have to protect the professors."

Controlling sex in advertising appears to have two principal directions. One is the control of advertising in such a way that specific audiences are protected from potentially harmful exposure. The other is the control of sexual content so that it conforms to certain standards, standards usually proposed by an individual or group of like-thinking individuals. I have sympathy with the goals of the former type of control, and considerable antipathy towards the latter type of control.

The protection of audiences from things they should not see is, as a concept, a highly desirable end. I think we are all aware, however, of the difficulty in defining what is and what is not appropriate for a given audience, and, once that is achieved, determining the extent to which dissemination to the public must be curtailed to ensure that the specific audience is not in peril of being exposed. As individuals, we should all have opinions on what these definitions are. I say we should have these opinions since the Supreme Court requires us to have them, given the current ruling on pornography which puts the responsibility for this definition in the hands of the relevant community.

But as researchers trying to discover ways of assisting the empirical definition of these things, we cannot afford to have opinions, indeed, we must protect ourselves against these opinions. The diligent researcher armed not only with hypotheses but with beliefs as well will generally prove his beliefs through his research efforts. One of the most eloquent statements of this fact was shown to me in the form of a declination by the Stanford Research Institute of an invitation to submit a research proposal concerning the determination of what educational philosophy was the most effective. The director of the Institute said, quite simply, that the research was impossible to conduct. Any effort to determine the answer in an objective sense would merely cloak the subjective proclivities of the researchers with scientific respectability.

We should not confuse our beliefs and our research outcomes; if we wish to persuade others of our point of view, then let us by all means become active and muster whatever support we can for our views. But we must not couple this with our research activities because our objectivity is gone and our results are worse than contaminated. Researchers should not be activists, and activists should not do research. Or at the very least, activists should have the common decency not to expect anyone else to listen to the research outcomes.

The second type of control, control of sexual content so that it conforms to certain standards, is to me as mean-

ingless an exercise as I can imagine. One of the sources of this opinion is related to some observations made in the course of living outside our country. In France I was struck by the extent of the use of unclothed females in advertising products in women's magazines. The illustrations were quite gorgeous, and I need not hide from you the fact that my readership of these magazines did not stem entirely from a desire to know what French womanhood was thinking. I am convinced that by whatever measure is employed, French women's magazines are much more explicit in their uncoverage of the female form, that the magazines are sexier than their US counterparts.

What makes this curious and also relevant to the idea of controlling the sex content of advertising is that in the US, we had had a research experience that suggested that such advertising would be found offensive and would not work. In Al Sawyer's dissertation research involving a laboratory experiment, he used ads in a large number of product categories, with several brands in each product category. In each case, one of the brands selected was chosen for its difference from the advertisements for the other brands. One of the product categories was ladies' foundation garments, and the "common" advertisements could be likened to the Sears catalog representation of the brands; stock photos of non-descript models wearing the products, about as arousing as a naked department-store manikin. The uncommon ad was for Bali bras, and was a very lovely and evocative photo, in soft focus and shot through textured glass in black and white, of a young lady, from the waist up apparently naked, demurely holding her arms in front of her breasts. When I say it was evocative, I do not mean provocative; Al and I admired the illustration and thought it was done in excellent taste, and there were no doubts about including it in the series. Compared to the dreary explicitness of the other brands' ads, we thought it would prove much more palatable over several repetitions than the other ads.

Making a long story slightly shorter, no other ad drew the wrath of our sample of housewives in a manner in any way comparable to the Bali ad. The spontaneous recall of the ad was very high, and was usually coupled with very derogatory comments about the sexual content of the ad. The study concerned itself with repetition effects, and the situation worsened as the repetitions increased; respondents tended to exaggerate considerably the number of times that they were shown that ad.

Based on these reactions, I feel free to conjecture that US housewives would find the advertising in French women's magazines highly offensive. And yet the French do not. Cultural relativism is, of course, not a new thing, but the point here is that control of sexual content in advertising is very much a culture-bound concept. To the extent to which we wish to mire our habits and attitudes in some specific mold and hold it constant, then we should be willing to control advertising's content, including its sexual content. To the extent that we want it to follow the modalities and mores of our culture, it should be set free. Advertising's content does not, will not, and cannot lead the shaping of our culture; it follows, a cautious step behind the avant garde, taking careful aim at the willingness of the great unwashed masses to sustain its content. And advertising is successful at this; why else should one-third of Bob Hope's and Johnny Carson's monologues be devoted to the quips about brands, products, and advertising?

If a more modern example of the cultural differences is desired, it is only necessary to note that in the US, lingerie manufacturers are prohibited from showing their wares on live models in TV commercials. Compare this to the French ad, run last year, in which a waiter enters a hotel room occupied by two young ladies to lay out their

room-service breakfast. The girls, clad only in panties, whisk in and out of the room as the hapless fellow tries to control his roving eyes, only to find them riveted on two eggs, sunnyside up, as one of the laughing girls finishes putting on her brassiere. The commercial is a thing of great beauty, humor, and taste; and it would not work at all in the US, even if live models could be used.

This is the source of my antipathy towards the suggestion that there is some way in which women "should" be portrayed in advertising. When our culture is ready for it, women in advertisements will all be liberated and successful; when our culture is ready for it, husbands will cease to bumble, the number of cowboys represented in advertisements will be proportional to the number of cowboys in the general population and, perhaps, Pete Rose will slug that sappy female reporter in the mouth and regain his manhood.

#### How (Sex in) Advertising Works

Earlier in this paper, I chided the Scigliimpaglia, Belch, and Cain paper and the Wilson and Moore paper for their eagerness in co-opting the methodology of the erotic laboratory. But these two papers do focus attention on the more important issues of how sexual content works in the context of advertising. The two papers provide a broad range of theoretical background and some empirical data which help to focus our attention on the design and execution of effective advertising utilizing sexual content. As the reader will by now have detected, in my own view the issue of sex in advertising is so broad that I feel that these papers and research efforts only begin to tell us what we would like to know and apparently cannot resist asking about sex in advertising.

In their reviews of the literature and their own empirical content, these papers clearly indicate that response to sex in advertising is not uniform across the population. Projecting from these findings to what I would expect will continue to be the outcome of researchers in sex in advertising, that is, more and more findings that subsets of the population based on sex and age of the respondents, other demographic factors such as education, and perhaps even psychological makeup, differ greatly from one another in their reaction to the use of specific sexual stimuli in advertising, I foresee the great problem in this area being that of drawing useful generalities.

Returning to earlier arguments, the problem that faces us is that sex is a primary part of the psychological baggage that we carry around with us. Our attitudes towards, abilities to cope with, and acceptance of sex play a dominant role in our willingness to acknowledge and receive messages in which sex plays an important, or even a peripheral part. This, of course, is the sort of situation in which sexual innuendo is a highly successful ploy: The message presented in such a way that those most likely to be offended have ample opportunity to selectively avoid reception of the offending portion of the message.

Here we arrive at the critical point of my argument. However it works, advertising is not a simple stimulus. Large, obscenely large, sums of money are spent to assure that the creative work on a campaign creates a richness of image and association well beyond that of a fuzzily reproduced Polaroid in the back pages of Hustler or a pre-Castro Cuban porno flick. To isolate, in this wealth of imagery and associations, those elements which arouse one person is a formidable task: To presume to do so for the population at large is foolhardy. Research on sex in advertising will lead us soon to the same point we have reached with respect to fear appeal

in advertising, although in my own belief fear is a much simpler phenomenon to deal with than sex. That is, that we will find that the sexier the message, the greater the effect, up to the point where anxiety begins to impair the reception of the message. And that segments can be identified whose anxieties are greater than others. We will be led to say, "Sex works, depending on the audience, the product, and so forth." This may sound as though I am very pessimistic about the chances of research in this area bringing us a lot of new insights in the foreseeable future. That possibility I leave up to you as, in the fatherly words of the Hawaiian cannibal, "One man's meat is another man's poi, son."

Two words of caution about null hypotheses seem in order. Expect differences. The greatest surprise for me would be the observation that a general truth had been tapped. Do not be satisfied with showing us differences significant at some level. Look for the unifying factors and the underlying generality. These are the things that will assist us in our understanding and our applications: these are the findings that would make our brand of beast superior to the cunning and conniving creative director who has long known what we have merely begun to suspect.

E. B. White said:

Humor can be dissected as a frog  
can, but the thing dies in the process  
and the innards are discouraging to  
any but the pure scientific mind.

Good things can come from the dissection process--there's a special providence in the fall of a frog--but those of us who love the living thing must register in this quiet way our small measure of dismay at the dissection of sex in advertising. We call for artistry and a sharp blade, fearing butchery.

Joel Saegert, The University of Texas at San Antonio

### Abstract

The levels-of-processing theory of memory maintains that stimulus material will be remembered as a function of the amount of personal elaboration to which it is subjected by the observer. An experiment was conducted which varied level of processing of brand names by directing attention to either physical or semantic aspects. Memory was superior for items presumed to be processed at a deep level.

### Introduction

Olson (1977) has recently discussed the potential of the "levels-of-processing" theory of memory (Craik and Lockhart, 1972) for understanding the processes which underlie advertising recall. This theory postulates that memory is a function of the type of cognitive processes which occur during exposure to a to-be-remembered stimulus and has elements in common with a number of recent theories of advertising effects (e.g., Leavitt, 1974; Wright, 1974).

Craik and Lockhart argued that memory is a function of the "level" to which material is processed and that this is independent of the amount of repetition or rehearsal of the stimulus. While a precise definition of "level" has not been specified, the variable involves the degree to which material is subjected to elaboration by the learner in terms of relating it to his prior experience and knowledge. Items processed to a deeper level (that is, subjected to greater elaboration) are hypothesized to be more readily remembered.

An example of a typical levels-of-processing experiment will serve to illustrate how the concept is operationally defined. Hyde and Jenkins (1973) manipulated the orienting task of a number of groups of experimental subjects. Each group was presented with a long list of words that were to be processed in one of several ways. In one condition, the subjects were asked to look at the words and rate how pleasant or unpleasant each one was on a five-point scale. Another group was simply asked to count and write down the number of times the letter E appeared in each word they were shown. Following this, the subjects in both groups were given a surprise recall task in which they were to write down all the words they could remember. This procedure corresponds to the usual "incidental learning" paradigm used in many memory experiments.

Differences in recall were observed among the groups. The letter counting group was only able to recall about 28% of the words while the pleasantness rating group got 48% correct. Remarkably, this was even better recall than that of a group of subjects who had been specifically instructed to try to learn the words.

It is presumed that the groups differed only in the level of processing of the material since the word list and exposure time were identical across conditions. In terms of Craik and Lockhart's theory, it is argued that when a person is required to ascertain whether a word is pleasant or unpleasant, it must be considered (thought about) in terms of all of its available attributes and associations; such deep processing is sufficient to maintain a high level of memorability. The letter-counting activity, on the other hand, directs the atten-

tion to superficial aspects of the material to be remembered, resulting in a shallow level of processing.

The levels theory has been subjected to a great deal of theoretical and empirical elaboration in the past several years. For example, Craik and Tulving (1975) have argued that the effects are more accurately characterized as "spread" of processing rather than depth. That is, the successful recall of an item depends upon its encoding in additional contexts (e.g., structural, phonemic or semantic) of the existing cognitive make-up of the individual.

In Craik and Tulving's experiments, subjects were asked to respond to questions about words presented to them. Again, no mention was made that the words were to be remembered. An example of a deep processing question would be "Is the word a part of the body?"; in this case, the attention is directed to the meaning (category membership) of the word. A shallow processing question example would be "Is the word in upper case letters?", while an intermediate processing question was "Does the word rhyme with weight?". A series of ten experiments demonstrated consistently better memory for deeper levels of processing. The experiments also ruled out the possibility that longer processing time for deep questions could account for the superior performance; this was done by asking questions which took a long time to answer but still only required shallow processing. Finally, Craik and Tulving varied the degree of elaboration required to respond to questions and found better recall for more elaborate processing.

The concept of memory maintenance as a function of the quality rather than the quantity of processing would seem to be of profound significance for information processing theories of consumer behavior. The present experiment was designed to extend Craik and Tulving's operational definition of depth of processing to an advertising context. The task required subjects to attend to brand names in advertisements; for half the ads, attention was directed to physical features while for the other half, semantic features were stressed. In this way, an attempt was made to manipulate the level at which each ad was processed.

### Methodology

#### Materials

Forty magazine advertisements were used as presentation stimuli. The ads were chosen to represent a wide variety of products which would vary considerably on such features as familiarity, distinctiveness or personal relevance. Examples included Pioneer Stereo, Dole Pineapple, Arrow Shirts and American Airlines. Subjects were asked to respond to a single question about each ad. The questions used to elicit deep processing concerned personal experience with the brands advertised; examples include "Do you have this brand in your home?" and "Have you ever bought this brand?". Shallow processing questions directed the attention of the observer to the formal aspects of the brand name; "Is the brand name in script letters?" and "Are there seven letters in the brand name?".

To insure that any observed difference in memory for the ads was not specific to the ads chosen for each condition, a control procedure was used. This involved shifting questions for each ad from the shallow to the deep condition or vice versa after half of the participants in the study had been tested. Thus, for half the subjects, one half of the ads (arbitrarily chosen) had shallow questions while for the other half of the subjects, that same half of the ads had deep processing questions. The reverse was true for the other half of the ads.

Brand names were randomly dispersed throughout the presentation list with the restriction that no more than three ads with questions of the same level of processing appeared contiguously.

#### Procedure

Subjects were seated at a table in a small room. The experiment was explained as an attempt to study perceptions of brand names in ads. The subjects were to look briefly at a number of ads and quickly answer a question about each one. An example was given of what was meant by "brand name" and a sample question was asked. The instructions indicated that a "yes" or "no" response was to be checked on a numbered answer sheet as soon as each question was asked.

The experimenter presented the ads by turning the pages of a notebook to the beat of an electronic timer set at a five-second rate. As each page was turned, the appropriate question was read. Subjects were given no prior indication that they would later be asked to remember the ads; post-experimental interviews indicated that the memory tests were indeed a surprise. "Buffer items", indistinguishable from other items but not scored, were included at the first and last of the list of ads to control for "serial position effects".

#### Dependent Variables

Two dependent variables, recall and recognition, were measured. In the recall test, the subjects were told to write down as many of the brand names as they could remember. This immediately followed presentation and two minutes were allowed for recall. For the recognition test, the subjects were given a list containing 80 brand names and were told to check "yes" or "no" to indicate whether they had seen each brand name in the stimulus notebook. Half of these items had actually been presented in the experiment while the other half were distractor items. The distractors were brand names in the same product class as the originally presented ads, but they had not been seen in the experiment. Thus, in the recognition test, each "old" brand name was "matched" with a "new" brand of the same product. Previously presented items and distractors were randomly dispersed through the list with the restriction that no more than three items of each type appear contiguously.

The recognition scores were corrected for guessing by subtracting the number of false alarms (brand names incorrectly checked "yes") from the number of hits (brand names correctly checked "yes"). The logic for this standard correction procedure is that each subject might be expected to guess "yes" by chance about as often for correct items as they would for incorrect ones.

#### Subjects

The subjects in the experiment were 30 adults (members of the secretarial staff) at a university in a southwestern metropolis. Fifteen subjects served in each group and approximately half were males.

## Results

The mean number of brand names correctly recalled and the mean number of brand names recognized (corrected for guessing) for the two types of questions for both subject groups are presented in Table 1. As can be seen, about twice as many brand names for which deep processing questions were asked were recalled compared

TABLE 1

MEAN BRAND NAME RECALL AND RECOGNITION  
(Standard deviations in parentheses)

		Type of Question	
		Shallow	Deep
Recall	Group 1	2.27 (1.91)	4.13 (2.03)
	Group 2	1.73 (1.28)	4.37 (1.99)
	Mean	2.50	4.75
Recognition	Group 1	9.00 (4.63)	13.00 (3.02)
	Group 2	8.60 (4.27)	13.73 (3.11)
	Mean	8.80	13.36

to those for which shallow processing questions were asked. (The "groups" variable refers to the control procedure of reversing deep and shallow questions for the first and second half of the subjects.) The deep processing question ads were also recognized substantially better.

Table 2 presents the summary tables for analyses of variance of the two dependent variables. The design

TABLE 2

ANALYSIS OF VARIANCE SUMMARY TABLES

Source		df	Mean Square	F	p
Recall	Group	1	.02	<1.0	ns
	Error between	28	3.27		
	Type of Question	1	84.02	24.57	<.001
	Interaction	1	3.75	1.10	ns
Recognition	Error within	28	3.42		
	Group	1	4.27	<1.0	ns
	Error between	28	21.48		
	Type of Question	1	365.07	47.32	<.001
	Interaction	1	13.07	1.69	ns
	Error within	28	7.71		

was a 2 x 2 factorial with type of question as a repeated measures variable. In both recall and recognition, the differences between the deep and shallow question items were significant ( $p < .001$ ); the small differences between groups and the interactions between groups and type of question were not significant.

Thus, for both recall and recognition, retention was superior for ads presented in the context of deep processing questions.

#### Discussion

While it cannot be claimed that the situation studied in the present experiment is precisely analogous to the experience of consumers viewing ads in the "real world", it can be said that there are elements in common. When a researcher calls to ask what ads were seen last night on TV, it is likely that the level to which the interviewee related ads to personal experience will influence recallability. In this context, it has been demonstrated in the present study that the recall and recognition of ads varies as a function of the type of questions asked and hence, the level of processing to which the ad was subjected.

Of course, the advertiser who wishes to elicit deep processing by the consumer cannot ask questions which will cause brand names to be related to a person's personal experience. However, ads can be designed in such a way that they elicit active participation by the viewer. That is, any ad which invites and effects the participation of the viewer by requiring him to supply some omitted aspect or to compare a brand with his personal experience or to predict the outcome of a hypothetical situation might be expected to require comparison of the brand name with previous knowledge. The result may be that the brand name becomes an integral part of the viewer's knowledge and hence is available for future retrieval.

Levels-of-processing theory places emphasis in the memory situation on the kind of processing that material receives rather than the number of exposures. While such a distinction between quality and quantity of processing may come as no surprise to advertising strategists, there has been an emphasis on the use of associative theory to explain repetition effects in advertising (e.g., Sawyer, 1974). Although it cannot be disputed that repetition is related to the likelihood that an ad will be remembered, it may be that repetition chiefly serves to provide multiple opportunities for observers to apply cognitive elaboration to the ad material, rather than to build up "habit strength" in accordance with classical associative theory. It seems safe to assume that such cognitive theories as Craik and Lockhart's levels of processing will receive a great deal of attention in future attempts to explain the memory of advertisements.

#### References

- F.I.M. Craik, and R.S. Lockhart, "Levels of Processing: A Framework for Memory Research," Journal of Verbal Learning and Verbal Behavior, 11 (1972), 671-684.
- F.I.M. Craik, and E. Tulving, "Depth of Processing and the Retention of Words in Episodic Memory," Journal of Experimental Psychology: General 104 (1975), 268-294.
- T.S. Hyde, and J.J. Jenkins, "Recall for Words as a Function of Semantic, Graphic and Syntactic Orienting Tasks," Journal of Verbal Learning and Verbal Behavior, 12 (1973), 471-480.
- C. Leavitt, "Strong versus Weak Effects of Mass Communication: Two Alternative Hypotheses," in G.D. Hughes and M.L. Ray, (Eds.) Buyer/Consumer Information Processing. (Chapel Hill: University of North Carolina Press, 1974, 256-269).

J.C. Olson, "Theories of Information Encoding and Storage: Implications for Consumer Research," Working Series in Marketing Research, No. 65 (August, 1977), College of Business Administration, The Pennsylvania State University.

A.G. Sawyer, "The Effects of Repetition: Conclusions and Suggestions about Experimental Laboratory Research," in G.D. Hughes and M.L. Ray, (Eds.) Buyer/Consumer Information Processing. (Chapel Hill: University of North Carolina Press, 1974, 190-219.)

P.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.)

PERCEPTUAL DISCREPANCIES IN THE TIME DURATION  
AND NUMBER OF TELEVISION COMMERCIALS

Peter H. Webb  
University of Illinois

Abstract

In recent years, increasing pressure has been applied to television networks to control both the quality and quantity of advertising. The industry has denied major increases in commercial time in the last 25 years. This research shows that people's perceptions of the amount of advertising on television are substantially different than the actual amounts, and also differ as a function of how the commercials are scheduled within programs.

Introduction

In recent years, debate about the contents of television has steadily increased. With television reaching deeper and deeper into the lives of more and more Americans (the set in the average television household, which includes 96% of all households, is now on more than six hours a day), a growing number of public interest groups, government agencies, program sponsors, and viewers themselves are expressing concern over the quality of this mass medium. Much of this concern focuses on aspects of programming, such as the amount of sex and violence reaching younger audiences or the limited amount of public service programming available to all types of audiences. However, non-program material, including commercials, promotional announcements, public service announcements, billboards, station breaks, and credits, has also received its share of attention. For example, cigarette commercials were banned from the airwaves in 1971, the amount of advertising on Saturday morning children's programming was limited in 1974, and the Federal Trade Commission is currently contemplating further restrictions on advertising to children.

The sheer volume of non-program material is a problem of concern for several groups. "Clutter", as labeled by broadcasting practitioners, is a problem for advertisers, due to potential decreases in effectiveness of commercial messages in a crowded environment. It is a problem for TV broadcasters, with increasing operating costs on the one hand and pressures from advertising agencies and government to reduce revenue generating non-program material on the other hand. Finally, it is a problem for viewers, who presumably turn on the set for the programming it offers, but all too often face frequent and lengthy program interruptions for commercials, promotional announcements and other non-program elements.

Although concern that viewers are being swamped with increasing amounts of advertising on television (advertising accounts for more than 80% of all non-program material) has resulted in threats of involuntary regulation from government agencies and threats of boycotts from advertisers and advertising agencies, the television industry has steadfastly maintained that the amount of advertising has not increased appreciably in the past 25 years. This raises several issues.

The first issue concerns how one defines the "amount" of advertising. Does this mean the amount of time devoted to advertising, the number of commercial interruptions, or the total number of commercials? In 1952, the Code Authority of the National Association of Broadcasters, a self-regulating agency of the industry, set the maximum allowable range for advertising from seven minutes per hour for prime time network television

to nine and one half minutes for non-prime time local television. Comparable figures in 1971 were ten minutes and sixteen minutes, increases of 43% and 68% respectively.<sup>1</sup>

With respect to commercial interruptions, restrictions on the maximum allowable number of interruptions by the Code Authority (a complex rule differing by time of day and type of program), has led to subtle changes in programming practices including the widespread use of "lead-ins" and "trailers", set apart from the main body of the program. The result is probably more interruptions than ever, although no longitudinal data is available.

Finally, although data do not extend as far back as the early days of television, results of an extensive study published in 1972 by the American Association of Advertising Agencies indicated a 50% increase in the number of commercials (and a 33% increase in the number of commercial minutes) over the eight year period from 1963 to 1971. Certainly one of the major contributions to the increase in the number of commercials has been the shift from the sixty second to the thirty second commercial.

Another important issue concerning the amount of advertising on television is viewer perceptions. Even if the actual amount of time devoted to advertising is neither "large" (at least relative to other media--another defense of current practices put forth by the industry) nor increasing substantially, what if viewers perceive this to be the case? Is this not a cause for concern? The viewer appears to have received little attention in this matter. Even in Steiner's 1960 landmark study, The People Look at Television (published in 1963), the one section of his lengthy questionnaire having to do with advertising focused almost entirely on content, with little attention on quantity. However, in an open-ended question concerning the irritating aspects of advertising, constant interruptions and frequent repetitions of the same commercials were common responses. A careful replication of this study ten years later showed similar, and in some cases even stronger results (Bower, 1973). An additional item in the replication study not included in Steiner's study revealed that 70% of the respondents felt television has too many commercials. Beyond these relatively sparse data, little exists to indicate exactly how people perceive the amount of television advertising (or other elements of non-program material).

As part of a larger study of clutter on television, data were gathered concerning just such perceptions. After exposure to videotaped television programming, respondents were asked a series of questions about the commercials including how many there were and how much time was devoted to them. The little that could be found in the literature to base any hypotheses on is presented in the next section. Following this, the specific studies are described and the results are presented.

<sup>1</sup>1978 allowances range from nine and a half to sixteen minutes.

## Time Perception

One aspect of cognition that has received little attention in advertising research is perception of time duration. This is particularly relevant to the problem of clutter since the amount of time devoted to non-program material is one way of operationally defining clutter. As noted earlier, practitioners frequently argue over whether or not clutter is actually increasing, but rarely present evidence of viewers' perceptions. Burke Marketing Research, Inc. (1972) did report an increase over a two-year period in people's perceptions of the total amount of television advertising. In extensive telephone interviews, Burke found that 67% of a nationwide sample of 1750 respondents perceived more advertising on television in 1972 than in 1970. In addition, 50% of the respondents felt that there was more advertising than programs were worth, while just 3% felt the opposite. No precise data on how much advertising had increased were ascertained.

The only other research relating directly to time perceptions of advertising was a 1966 study of "piggyback" commercials by Martilla and Thompson. Piggybacking refers to the practice of allocating a commercial time slot to advertisements for two or more unrelated products. In this study, test commercials included 60 and 30 second versions of five commercials. For the 30 second version, a second 30 second commercial was "piggybacked" either in front of or behind the test commercial. Commercials were embedded in one of two 30 minute programs, which respondents watched in a laboratory setting.

When questioned immediately following the presentations, on the average the 60 second test commercial was judged to be 67 seconds long (a control 60 second commercial in a different position in the program was estimated on the average as 64 seconds long). However, the two 30 second commercials appearing in the same position as the 60 received a combined estimate of 83 seconds—a 38% overestimate of the actual time compared with a 12% overestimate for the 60 second version.

Although this finding confirmed Martilla and Thompson's hypothesis that time perception of advertisements increases as the number of advertisements increases even if the actual total time remains constant, no explanation was offered for this difference. What factors might relate to consumer misperceptions of time?

Unfortunately, most behavioral research involving time treats it as a fixed parameter and not a dependent variable. But there are some studies in which time estimates have been measured as a function of varying physical situations or task characteristics (Arvedson, 1974). Two such studies suggest several dimensions relevant to clutter which may affect estimates of time duration.

Ornstein (1968) had subjects study a set of drawings of varying complexity. He then asked subjects to estimate how long they had been given to study the drawings. He found that the more complex the drawing the longer the time interval was perceived to be (in actuality, the interval was 30 seconds in every case). This study was replicated by Hogan (1975), using slightly different stimuli, but varying the same dimension. The pictures were presented for just fifteen seconds in his experiment. Unlike Ornstein, Hogan found a curvilinear relationship between complexity and perceived time duration. The lowest estimates were given for moderately complex stimuli. Higher estimates were made for the simpler and the more complex drawings.

Although these two studies show seemingly different results (it may be that Ornstein did not include stimuli

simple enough to show the curvilinear relationship found by Hogan), they do suggest that a systematic relationship does exist between stimulus complexity and perceived duration of time. To the extent that continuity is related to complexity, a string of unrelated non-program elements is likely to be judged more complex than an integrated program segment. This may cause an overestimate of the time devoted to non-program material. Thus from a time duration standpoint, people may judge clutter to be worse than it actually is. This contention is supported by Martilla and Thompson's (1966) finding that two 30 second commercials were perceived as longer than a single 60 second commercial.

It should be noted that at the time this study was done in 1966, 30 second commercials were not nearly as dominant as they are today. It may be that the overestimate in duration for the two 30's was due to a variance between actual scheduling and viewers' expectations of scheduling. In fact, although there is no direct evidence, it may be speculated that viewer expectation lies at the heart of many reactions to clutter, including time perception. The overall hypothesis would be that the closer the match between actual and anticipated characteristics of an advertisement or a group of advertisements, the more accurate would be the perceived time duration. Variance from expectation of any kind may lead to distorted estimates of time perception.

Both Ornstein's and Hogan's studies indicate distortions are likely to be positive—that overestimates of time duration are far more common than underestimates. Thus variation in time perception may occur not only as a function of complexity, but also as a function of other variables such as relevance, affect, and novelty (both absolute and relative to surrounding material). For example, an advertisement seen as more or less relevant than expected would be perceived as longer than one whose degree of relevance matched expectations. The same phenomenon would occur for advertisements seen as more pleasant, more unpleasant, or about as pleasant as expected. Finally, changes in scheduling patterns for non-program material would create divergences from expectations that could alter perceptions of duration of elements comprising the schedules.

## Study One

In the first study of television clutter, subjects were brought to a central location to view closed-circuit programs. The subjects were told that they were participating in a study of humor and violence on television. Participants saw two program segments (each approximately ten minutes in length) taped directly off the air at an earlier date: one from a prime time situation comedy, the other from a prime time crime drama series. Each segment included 30 seconds of lead-in from the beginning of the show and was edited in such a way that it could "stand by itself" as representing a complete story.

Placed between the two program segments was a simulated station break that contained various elements of clutter comprising the experimental treatments. Thus the videotaped elements employed in this study were edited and combined in such a way as to realistically represent the transition from one television program to another. Subjects saw either zero, four, or eight commercials, credits for the first program, and a station identification. All commercials were 30 seconds in length. Specifically, the two tapes representing "low" clutter conditions contained four commercials in two groups of two each separated by the program credits (30 seconds) and a five second station identification. The two tapes representing "high" clutter conditions contained eight commercials (including two public service announcements) in



two groups of four each, again separated by the program credits and station ID. A fifth tape contained the credits and station ID, but no commercials at all.

A total of 192 subjects viewed the program, both men and women aged 18 to 55, from predominantly middle income families in suburban San Francisco. Each of the five tapes was viewed by eight groups of five subjects (eight of the forty groups were missing one subject). Subjects viewed the tapes in a room set up to look like a living room. They were instructed to act as they would watching TV at home, including talking with each other if they wished, getting coffee and cookies (placed at the back of the room), or browsing through magazines placed on a coffee table in front of the TV monitor.

Following the presentations, subjects filled out a questionnaire that included a series of questions about the programs (designed to be consistent with the cover story and to eliminate short term memory traces), questions about the contents of commercials, and questions about the number and duration of commercials. The latter questions were simply, "How many commercials were there?" and "How much time was devoted to commercials?"

**Results**

Table 1 shows estimates of number and time devoted to commercials for high and low clutter conditions compared

TABLE 1

COMMERCIAL TIME AND NUMBER ESTIMATES BY CLUTTER LEVEL

Clutter Level	N	Average Perceived Time (Minutes)	Actual Time (Minutes)	Average Over-Estimate (Minutes)
Baseline	(32)	1.46	0	1.46*
Low	(77)	4.65	2.00	2.65***
High	(71)	5.87	4.12 <sup>1</sup>	1.75***

Clutter Level	N	Average Perceived Number	Actual Number	Average Over (under)-Estimate
Baseline	(37)	1.05	0	1.05**
Low	(76)	4.15	4	0.15
High	(73)	7.21	8	(0.79)

\*p < .02  
 \*\*p < .01  
 \*\*\*p < .001

<sup>1</sup>Anti-drug public service announcement inaccurately edited from 60 seconds to 37 seconds instead of 30 seconds.

with actual amounts. These figures are raw estimates. Originally, it was thought that these estimates should be "normalized" by perceived length or presentations to eliminate errors in estimation due only to poor sense of time duration. However, 46 percent of the subjects estimated the presentations to be 30 minutes long. Since this was probably due to expectation rather than actual perception, the adjusted measure of commercial duration was dropped. As hypothesized, the time devoted to commercials was significantly overestimated in all conditions. In low clutter (four commercials), 81% of the

subjects overestimated commercial time, compared with 65% in high clutter (eight commercials). The average amounts of the overestimates were 2.65 minutes for low clutter (t = 8.33, p < .001) and 1.75 minutes for high clutter (t = 4.96, p < .001). Even when no commercials were aired, the average subject perceived that the presentation contained almost a minute and a half of commercials. This is probably attributable to prior experience with television programming. Even when commercials cannot be recalled, people assume they were present since this is the case with all programming except on public television.

The influence of prior experience is also evident from the degree of over-estimation in high and low clutter conditions. When the amount of commercial time is increased by a factor of two, the estimated amount of time increases by just 26%. In the absence of prior experience, this percentage would probably be much greater.

With respect to number of commercials, estimates were surprisingly accurate. Subjects in the low clutter conditions (four commercials) perceived an average of 4.15 commercials. Subjects in the high clutter conditions (eight commercials) perceived an average of 7.21 commercials. In the baseline condition (zero commercials), subjects perceived an average of 1.05 (the only significant difference from actuality: t = 2.55, p < .01). Thus when few or no commercials were aired, subjects over-estimated the number, but when a large number of commercials were aired, subjects under-estimated the number. Once again, prior experience appears to have guided perceptions. Actual station breaks, such as the one simulated in this study, usually contain fewer than eight, but more than two or three commercials. It is interesting to note that deviations in perceived number of commercials are much smaller than those in amount of time devoted to commercials.

It is also interesting to note that subjects perceived the average commercial to be significantly longer than 30 seconds. Dividing their estimates of time devoted to commercials by estimates of the number of commercials shows an average perceived commercial duration of 49 seconds in high clutter, 67 seconds in low clutter, and 83 seconds in the baseline condition. This, of course, is an indirect estimate of commercial length, and might differ from explicit estimates of length which were not included in the questionnaire.

Study Two

The second study of clutter focused on alternative scheduling patterns of commercials within programs. Once again, subjects were brought to a central location to view closed-circuit television programs. With a few minor alterations, the methodology employed was identical to that of the first study. Having been told they were participating in a study of television ratings, subjects viewed one of two half-hour network programs: a prime time situation comedy or a daytime serial. Twelve 30 second commercials were edited into the programs in one of three different schedules: Six breaks containing two commercials each (6 X 2), three breaks containing four commercials each (3 X 4); or two breaks containing six commercials each (2 X 6).

The 6 X 2 condition is typical of commercial scheduling patterns actually used for daytime serials. The 3 X 4 condition is a typical interruption pattern for prime time programming, although in this instance, the number of commercials was greater than would ordinarily be the case in prime time. This was due to the desire to hold the number of commercials and the time devoted to them constant in order to focus on differences due only to the scheduling patterns or the programs. The 2 X 6

condition is atypical of any current practices, but is interesting in view of previous findings that people would prefer a "bunching" of commercials into fewer interruptions (Burke, 1972).

A total of 189 subjects viewed the programs. As in the first study, the subjects were aged 18 to 55 from mostly middle income families in suburban San Francisco. In this study, subjects were all women. Nobody participated in both studies. Each of the six tapes (two programs; three commercial schedules) was viewed by eight groups of four subjects (three of the 48 groups were missing one subject). In all other respects, the study was conducted in exactly the same manner as the first study.

#### Results

Table 2 shows estimates of number and time devoted to commercials for the different programming and commercial scheduling conditions.<sup>2</sup> It is interesting to

TABLE 2  
COMMERCIAL TIME AND NUMBER ESTIMATES BY  
PROGRAM AND COMMERCIAL SCHEDULE

Program	N	Average Time Estimate (Minutes)	Actual Amount (Minutes)
Prime time	(92)	9.94	6
Daytime	(95)	9.76	6
<u>Commercial Schedule</u>			
6 X 2	(63)	11.00	6
3 X 4	(59)	9.48	6
2 X 6	(63)	9.05	6
Program	N	Average Number of Commercials (Estimates)	Actual Number of Commercials
Prime time	(92)	9.24	12
Daytime	(95)	9.16	12
<u>Commercial Schedule</u>			
6 X 2	(63)	9.25	12
3 X 4	(61)	8.82	12
2 X 6	(63)	9.51	12

note that for both programs and for all three scheduling

<sup>2</sup>As was the case in study one, a high percentage of subjects estimated the presentation to be 30 minutes long (48%). This was probably due more to expectation than actual perception. As a result, time perception data for commercials were not weighted by presentation length estimates as originally intended.

conditions, the amount of time devoted to commercials was significantly overestimated (by as much as 83%). However, the perceived number of commercials was substantially underestimated. In fact, only 25 percent of the subjects perceived as many as twelve commercials--the actual number. This was the same pattern found in the first study. The amount of time devoted to commercials was consistently overestimated as hypothesized. However, the number of commercials was underestimated in the high clutter conditions (7.21, compared with eight in actuality). It can be seen from Table 2 that as the actual number of commercials further increases, the degree of underestimation also increases (23%, compared with 10% in the first study).

Based on the time duration findings of Ornstein and Hogan discussed earlier, it was hypothesized for this study that the perceived time devoted to commercials would increase as the number of interruptions increased. This hypothesis is based on the notion that increasing the complexity of stimuli within a fixed duration of time leads to increased estimates of that duration. It was here postulated that more interruptions produce a more complex stimulus "package" (program plus commercials). In addition, more interruptions are likely to be more annoying which also figures to increase the estimate of time duration.

From Table 2, the hypothesis is supported (from a one way analysis of variance,  $F = 4.05$ , 2 and 184 d.f.,  $p < .02$ ). The three interruption conditions were perceived as containing 5% more minutes of commercials than the two interruption conditions and the six interruption conditions 22% more. The estimated number of commercials does not reflect this pattern. The three break conditions show the minimum estimate, while the two break conditions actually show the maximum estimate, despite the minimum time estimate for these conditions. This may have occurred as a result of the two break conditions containing six consecutive commercials during each break, an unusually large number without at least a station break in between them.

It was also hypothesized that perceptions of commercial amount and duration would be higher for the prime time program, than for the daytime program. This was based on the fact that the number and duration of commercials included in the presentations was normal for daytime programming, but abnormally high for prime time programming. It can be seen from the table that the hypothesis is supported, although the differences are very small and not significant.

#### Summary and Implications

These studies of television clutter, with specific focus on perceptions of amount of advertising, show that wide variations exist between actuality and perceptions. The first of the two studies showed significant overestimates in the amount of time devoted to commercials. The second study showed significant variations in commercial time estimates as a function of the number of commercial breaks and how the breaks are positioned within programs.

These findings present potential problems, both for advertisers and for the television industry. Voluntary (NAB TV code subscription) or even federal regulation of the amount of non-program material on television based solely on actuality may not properly reflect viewer reaction. Thus, for example, as commercials get shorter, or as the number of commercial interruptions increases, even if the actual amount of advertising remains within established guidelines, viewers perceive such changes as representing substantial increases in the amount of advertising on television. This may produce negative feelings toward commercials and even

specific brands involved. Alternatively, it may result in actual decreases in television audiences. Evidence that this may already be occurring comes from Nielsen ratings indicating an "unexplained" drop in several daytime audiences in late 1977 (Advertising Age, 11/28/77).

Of a more positive nature, these studies suggest that negative effects on viewers or non-program material may be reduced by paying more attention to the viewer's perceptions of specific practices. Thus, for example, bunching of commercials into fewer interruptions may decrease the intrusiveness of the interruptions, resulting in more favorable reactions or even larger audiences. It should be cautioned, however, that such changes should be based on aspects of commercial response other than simply perceptions of number and duration. Traditional measures of response such as recall, attitudes and purchase intentions may reveal an inverse relationship between time perception as dealt with in these studies and commercial effectiveness.

#### References

"Another TV Mystery: Case of the Missing Young Women." Advertising Age 48 (November 28, 1977): 6.

Lennart A. Arvedson, "Deadlines and Organizational Behavior: A Laboratory Investigation of the Effects on Individual Task Performance." Unpublished doctoral dissertation, Stanford University, 1974.

Robert T. Bower, Television and the Public. New York: Holt, Rinehart and Winston, 1973.

Burke Marketing Research, Inc. "Viewer Attitudes Toward Commercial Clutter on Television and Media Buying Implications." Presentation to the 18th ARF Conference, November 14, 1972.

H. Wayne Hogan, "Time Perception and Stimulus Preference as a Function of Stimulus Complexity." Journal of Personality and Social Psychology, 31 (January, 1975): 32-35.

John A. Martilla and Donald L. Thompson, "The Perceived Effects of Piggyback Television Commercials." Journal of Marketing Research, 3 (November, 1966): 365-71.

Robert E. Ornstein, "On the Experience of Duration." Unpublished doctoral dissertation, Stanford University, 1968.

Gary B. Steiner, The People Look at Television. New York: Alfred A. Knopf, 1963.

Edwin C. Hackleman, The University of Connecticut  
Subhash C. Jain, The University of Connecticut

### Abstract

The recent surge of comparison advertising has raised a controversy about the effectiveness of such ads. This article reports the findings of an experiment conducted to measure the attitudes toward comparison ads vis-a-vis individual ads. Overall, subjects' attitudes toward comparison ads did not differ significantly from toward individual ads. Comparison ads, however, were found to be more effective in the case of shopping goods.

### Introduction

There is a growing trend in TV advertising today toward the comparison of one brand against one or more competitive brands through explicitly naming them on a variety of specific product or service attributes (Shimp, 1975). Such advertising, usually referred to as comparison advertising, began in early 1970's (Ulanoff, 1975). While there have been occasional comparison ads in the past, they were not in as explicit terms as we find them currently (Wilkie and Farris, 1975). One notices comparison ads being aired on all forms of goods and services.

There are no laws in the United States which prohibit comparison advertising. Traditionally, however, industry self-regulations discouraged such a practice. An important reason for the current surge in comparison advertising has been the Federal Trade Commission's encouragement for such advertising. The FTC did so in the hope that this would provide consumers adequate information to make buying decisions (Dougherty, 1973). Among the networks, the National Broadcasting Company was the first one to issue guidelines and accept comparison ads (Christopher, 1975). Since then both ABC and CBS have started doing so, too.

Needless to say, comparison advertising, being a new practice, has become a controversial issue among advertisers, media, advertising agencies and government. As an example, reacting on TWA's comparison ad, an airline executive called the TWA campaign flagrant, neither good for the industry nor the public (Hugh, 1976). According to an Ogilvy and Mather (a New York ad agency) study:

Comparative TV commercials offer no advantage to package goods advertisers in persuasion or brand identification. In fact, they reduce the believability of claims, make consumers more aware of competitors, and add confusion to the message (Advertising Age, May, 1976).

However, not all advertising practitioners decry comparison advertising. An ad agency executive remarked that it was an effective device to create consumer confidence in a brand (Advertising Age, May, 1976). Specifically, comparison advertising has been credited for increasing Schick's Fleximatic electric shaver market share from 8 percent to 24 percent (Business Week, May, 1975).

The above controversy has raised a furor among the advertisers. Over the past five years, according to an estimate of National Advertising Division of the Council of Better Business Bureaus, complaints involving comparative advertising have jumped to 38 percent of all complaints. The number of complaints filed by advertisers against competitors have risen five-fold since the advent of comparison advertising (Broadcasting, May, 1976).

Comparison advertising is an important form of advertising with legal, ethical, consumerism and promotional decision implications. To put comparison advertising in proper perspective, empirical and systematic research is required. Wilkie and Farris, drawing on the concepts of behavioral sciences, concluded that comparison advertising was a powerful tool for the marketers (Wilkie and Farris, 1975). Prasad's laboratory research showed an enhancement in message recall via comparison advertising (Prasad, 1976). Clarke's empirical study indicated that advertising influences not only the sales of the brand advertised, but also the sales of other brands (Clarke, 1973). Based on this study, comparison advertising should help in sales gains for all the brands included in the ad. While these studies have been interesting and revealing, they hardly suffice to resolve the controversy surrounding comparison advertising. Apparently, comparison advertising requires further empirical insights.

This article focuses on the effectiveness of comparison advertising from the viewpoint of consumers. It is based on an empirical study of a sample of male and female subjects on their attitude of comparison advertisements for different products. The hypotheses of the research are directed not only from previous work in the comparison advertising area, but also from a desire to uncover relationships which have not been investigated between strategic marketing mix variables. Specifically, the following statements are set forth:

1. Comparison advertisements will be more effective than non-comparison advertisements for some products but not for others.
2. Comparison advertisements will be no more effective for females than for males; also, non-comparison advertisements will be no more effective for females than for males.

The first hypothesis is not stated in a null sense because of the difference in buyer behavior and motivations surrounding the purchase of different products. Shopping goods, for example, constitute a much more salient purchase for the buyer and usually entail product-brand comparisons of style, quality, functions, price, and so on. It would seem, therefore, that comparison advertising would be more beneficial for shopping goods than convenience products in which shopping is much more routine. The second hypothesis is stated in the null sense for lack of empirical evidence to the contrary. However, life style differences between males and females could produce different views on comparison advertising messages. The implications for strategic marketing are sufficiently strong to justify investigation.

### Methodology

The hypotheses suggest the use of a split-plot repeated measures experimental design in which a subject receives all levels of some treatments but one level of other treatments. Myers (1969) and Winer (1971) refer to this layout as a mixed design or as a multifactor experiment having repeated measures on several elements. As Kirk (1969) emphasizes, the design is especially appropriate when the experimenter wishes to control a major source of subject heterogeneity -- a common problem in behavioral research.

Figure 1 illustrates the block diagram corresponding to the split-plot factorial 2.2 (12) experiment which was employed. It was desired to block the subjects by sex to isolate this source of variation from the other two major independent variables, product type and comparison/non-comparison advertising. The split-plot factorial design permits isolating block effects, main effects, and interacting effects -- all of which were considered possible influences in the variances of the dependent variable, attitude toward advertising.

Three categories of consumer goods were chosen: convenience, shopping, and specialty goods. Four products from each of these categories commonly used by both males and females were selected. For each of the 12 products, two brand names were prepared. The procedure used for the selection of brand names was much the same as that employed by Misra and Jain (1971). Briefly, 18 adult males and 18 adult females randomly selected from the list of charge customers of a major department store were asked to give within a 40-second period as many relevant associations as they could to each of the 12 products. For each product, a separate sheet of paper was provided. The subjects were instructed not to give any existing brand names that they might be aware of as their associative responses. The most frequent (occurring 15 times or more) responses were then determined. From these response words, four brand names were constructed. An additional randomly selected 40 adults, 20 males and 20 females, were given a list of 48 brand names (for 12 products) in random order without any mention of the product they might represent. These

subjects were asked to rate each brand name on a 5-point Likert scale with ends labeled "very meaningful" and "meaningless." The mean meaningfulness value for each brand name was calculated from these ratings following which two brand names with highest meaningfulness value were selected. These names are shown in Table 1

Three advertisements were found from magazines and newspapers for each of the twelve products. An effort was made to insure that the advertisements appeared similar, and all "real" brand names were removed from the ads. Of the three ads for a given product type, the first one was given the brand name shown in the first column of Table 1. The other two ads were combined to make one ad from both. Thus the resulting second ad was a comparison ad, i.e., it possessed not only the brand name in column one of Table 1, but also the brand name shown in the second column. In 6 comparison ads, the brand name in column one of Table 1 appeared first, while in the remaining 6 ads, the brand name in column two appeared first. Extensive care was taken to be sure that the individual and comparison ads for a brand were as similar as possible and possessing identical claims concerning product performance. Thus in all, there were 24 ads for 12 products. All the ads were tested on a 6-point realism scale by a group of 5 male and 5 female advertisers. Three of the 24 ads which scored below the mid-point were replaced using the procedure explained above.

Colored transparencies were made for all the 24 ads. These transparencies were shown to 120 subjects, 60 males and 60 females. The 120 subjects represented

FIGURE 1

BLOCK DIAGRAM OF THE SPLIT-PLOT FACTORIAL, 2.2(12) DESIGN

		$B_1$	$B_i$	$B_2$	$B_2 \cdot \cdot \cdot B_{12}$	$B_{12}$
		$C_1$	$C_2$	$C_1$	$C_2 \cdot \cdot \cdot C_1$	$C_2$
$A_1$	$S_1$	$Y_{1111}$	$Y_{1121}$	$Y_{1211}$	$Y_{1221} \cdot \cdot \cdot Y_{1(12)11}$	$Y_{1(12)21}$
	$S_2$	$Y_{1112}$	$Y_{1122}$	$Y_{1212}$	$Y_{1222} \cdot \cdot \cdot Y_{1(12)12}$	$Y_{1(12)22}$
	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$
	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$
	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$
	$S_{60}$	$Y_{111(60)}$	$Y_{112(60)}$	$Y_{121(60)}$	$Y_{112(60)} \cdot \cdot \cdot Y_{1(12)1(60)}$	$Y_{1(12)2(60)}$
$A_2$	$S_{16}$	$Y_{211(61)}$	$Y_{212(61)}$	$Y_{221(61)}$	$Y_{222(61)} \cdot \cdot \cdot Y_{2(12)1(61)}$	$Y_{2(12)2(61)}$
	$S_{17}$	$Y_{211(62)}$	$Y_{212(62)}$	$Y_{221(62)}$	$Y_{222(62)} \cdot \cdot \cdot Y_{2(12)1(62)}$	$Y_{2(12)2(62)}$
	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$
	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$
	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$
	$S_{120}$	$Y_{211(120)}$	$Y_{212(120)}$	$Y_{221(120)}$	$Y_{222(120)} \cdot \cdot \cdot Y_{2(12)1(120)}$	$Y_{2(12)2(120)}$

2 levels of  $A_i$  = sex  
 12 levels of  $B_j$  = product type  
 2 levels of  $C_k$  = advertising type  
 120 levels of  $S_m$  = subjects  
 2880 observations of  $Y_{ijklm}$  = attitude measurement

TABLE 1

## PRODUCTS AND BRAND NAMES INVOLVED IN THE ADVERTISEMENTS

	Brand I	Brand II
<u>Convenience Goods</u>		
1. Coffee	Brew Rich	E Z Perk
2. Toothpaste	Sprint	Sparkie
3. Headache Pill	Feel Good	Easyway
4. Cigaret	Smoky	Puffy
<u>Shopping Goods</u>		
5. Refrigerator	Beaut	Kitchen Style
6. Clock Radio	Accurate	Sleepwell
7. Camera	Take-a-Picture	Fantastic
8. Electric Shaver	Shave Well	Smoothy
<u>Specialty Goods</u>		
9. Piano	Musica	Pianoplayer
10. Foreign Sports Car	Space & Speed	Champion
11. Overseas Vacation	Duke & Duchess	Elite Fun
12. Foreign Liqueur	Soma	Majestic

three subsamples from three major population centers in the midwestern and eastern regions of the nation. The experiment, therefore, was replicated three times during shopping hours in the community rooms of large shopping centers located near the population centers. Each subsample was further subdivided into three subgroups of ten subjects each for transparency presentation (See Table 2). A counterbalancing approach was implemented for each product type, and the advertisements themselves were shown in a Latin square type arrangement in an effort to minimize order effects. The 12 x 12 Latin square configuration actually represents nine completely different 4 x 4 Latin squares, blocked in a 3 x 3 Latin square arrangement. With this design, no two subgroups received the same order of presentation of the 12 pairs of ads, and each subsample received a different order of presentation of the product types, (convenience, shopping, and specialty) features in the advertisements. In addition, to minimize any order effects due to the type of advertisement, five of the members of each subgroup saw the comparison ads first, whereas the other five saw the individual ads first. The entire presentation was terminated by a transparency that showed the name of a fictitious advertising agency.

The subjects were given four minutes to evaluate each advertisement on ten 7-point semantic differential scales. Each ad was coded, and the subject was instructed to copy the code number in a booklet as the ad was rated. Immediately following the ratings of all the 24 ads, the booklets were collected. Since the experiment was conducted to measure the effectiveness of the ads whose brand names appear in column one of Table 1, the subjects were told the name of the brand in which the advertising agency was interested.

The bipolar adjective pairs were selected by first investigating an instrument composed of 35 scales which had been commonly used in previous studies evaluating advertising messages. Three of the ads from Table 1 were randomly selected and evaluated using this instrument by 90 university-level students. The data from this pre-test was subjected to principal-components factor analysis. From these results, the ten scales exhibiting the best reliability and construct validity were selected for use in the shopping center experiment. These scales were pleasant-unpleasant, good-bad,

TABLE 2

## ORDER OF PRESENTATION OF ADVERTISING TRANSPARENCIES TO SUBSAMPLES AND SUBGROUPS

		Product Types											
Subsample	Subgroup	Convenience			Shopping			Specialty					
I	1	1	4	2	3	5	8	6	7	9	10	11	12
	2	4	1	3	2	6	7	8	5	12	11	10	9
	3	3	2	1	4	8	5	7	6	11	9	12	10
	4	2	3	4	1	7	6	5	8	10	12	9	11
II	5	Specialty			Convenience			Shopping					
	6	11	9	12	10	2	3	4	1	6	7	8	5
	7	12	11	10	9	3	2	1	4	5	8	6	7
	8	9	10	11	12	4	1	3	2	7	6	5	8
III	9	Shopping			Specialty			Convenience					
	10	8	5	7	6	10	12	9	11	4	1	3	2
	11	5	8	6	7	11	9	12	10	3	2	1	4
	12	6	7	8	5	9	10	11	12	2	3	4	1

ambiguous-clear, negative-positive, approving-disapproving, unimportant-important, unappealing-appealing, sincere-insincere, dynamic-dull, and depressing-refreshing. The order of presentation of the ten scales was randomized for each answer sheet, and for each scale the positions of polar adjectives were reversed half the time.

## Results

Table 3 reports the three-way analysis of variance performed on the observations in accordance with the design. With reference to the univariate F tests and their significance levels, the following findings concerning the original hypotheses are now presented:

TABLE 3

## ANALYSIS OF VARIANCE SOURCE TABLE

Source	SS	dF	MS	F
1. Between Subjects	105.90	119	.890	
2. A (Sex)	6.57	1	6.57	9.57*
3. Subj. W. Groups	81.04	118	.687	
4. Within Subjects	3371.53	2760	1.222	
5. B (Product)	412.64	11	37.51	38.79*
6. AB	37.47	11	3.41	4.38*
7. B x Subj. W. Groups	1255.68	1298	.967	
8. C (Advertisement)	.615	1	.615	.935
9. AC	.724	1	.724	1.10
10. C x Subj. W. Groups	77.65	118	.658	
11. BC	89.78	11	8.16	7.13*
12. ABC	11.75	11	1.068	.934
13. BC x Subj. W. Groups	1485.22	1298	1.144	
14. Total	3477.42	2879		

\* Significant at <.01 level.

1. The sex of an individual significantly affects attitude toward advertisements (<.01 level).
2. Attitudes toward advertisements are affected significantly by the product type being advertised (<.01 level).
3. There is no significant difference in attitude toward comparison ads and non-comparison ads.
4. Attitudes toward ads are significantly affected by the interaction between sex and product type (<.01 level).

5. Attitudes toward advertisements are affected significantly by the interaction between product type and type of advertisement (<.01 level).
6. The sex of an individual does not influence attitudes toward comparison or non-comparison advertisements.

A Scheffe (1953) test for pairwise multiple comparisons between treatment and cell means was performed, and results are summarized in Tables 4 and 5. Table 4 lists the mean attitude scores for the two significant

TABLE 4

MEAN ATTITUDE SCORES: SEX AND PRODUCT TYPE (MAIN EFFECTS)

Sex	Mean Score	Rank
Male (A <sub>1</sub> )	4.24 <sup>b</sup>	2
Female (A <sub>2</sub> )	4.44 <sup>a</sup>	1

Product Type	Mean Score	Rank
Coffee	3.82 <sup>b</sup>	10
Toothpaste	4.97 <sup>a</sup>	3
Headache Pill	4.82 <sup>a</sup>	4
Cigarettes	3.89 <sup>b</sup>	9
Refrigerator	4.52	5
Clock Radio	4.05	8
Camera	4.22	6
Electric Shaver	5.26 <sup>a</sup>	2
Piano	5.50 <sup>a</sup>	1
Foreign Sports Car	3.34 <sup>b</sup>	12
Overseas Vacation	4.16	7
Foreign Liqueur	3.55 <sup>b</sup>	11

Note: The higher the score, the more favorable the attitude toward the advertisement.

a - Significantly more favorable attitude than b at P <.05.

TABLE 5

MEAN ATTITUDE SCORES: PRODUCT TYPE BY SEX, PRODUCT TYPE BY ADVERTISING TYPE (INTERACTIONS)

Product Type	Sex		Advertising Message	
	Male	Female	Non-Comparison	Comparison
Coffee	3.74	3.90	4.13*	3.51
Toothpaste	4.94	5.00	4.79	5.15
Headache Pill	4.93	4.70	4.86	4.76
Cigarettes	3.53	4.25*	3.93	3.85**
Refrigerator	4.24	4.80*	3.92	5.12**
Clock Radio	4.25*	3.85	3.56*	4.54
Camera	4.14	4.30	4.60*	3.84
Electric Shaver	5.12	5.40	5.06	5.46*
Piano	5.50	5.50	5.70*	5.29
Foreign Sports Car	3.25	3.45	3.21	3.49
Overseas Vacation	3.96	4.35*	4.45*	3.87
Foreign Liqueur	3.35	3.75*	3.61	3.49

Note: The higher the score, the more favorable the attitude

\* Significantly higher at P <.05.

\*\* Significantly higher at P <.01.

main effects, sex and product type. Females held more favorable attitudes toward all advertising messages and for all product types in general than did males. However, there was no difference in male and female

attitudes toward either comparison or non-comparison ads overall. Hypothesis H<sub>2</sub>, therefore, is rejected.

Those products receiving the highest attitude scores for their ads were a piano, an electric shaver, toothpaste, and a headache pill. Ads for these products were rated significantly higher than ads for a foreign sports car, a foreign liqueur, coffee, and cigarettes. There appeared to be no evidence of a convenience-shopping-specialty grouping of goods which would allow one to conclude that ratings of one classification were higher or lower than another (without considering differences in advertising type).

Table 5 contains the mean attitude scores corresponding to the significant interactive effects, sex by product type and type of advertising message by product type. Females were more favorable in their attitudes toward ads for cigarettes, a refrigerator, an overseas vacation, and foreign liqueur. Males evaluated only the clock radio ad more favorably than the females. Table 5 also illustrates that a non-comparison ad for coffee, a camera, a piano, and an overseas vacation were evaluated more favorably than a comparison ad for the same products. On the other hand, a comparison ad was deemed more favorable than a non-comparison ad for a refrigerator, a clock radio, and an electric shaver. Since all three of these products are shopping goods, one is tempted to conclude that shopping goods do benefit from comparison advertising, whereas convenience and specialty goods tend to benefit more from non-comparison ads. This finding empirically supports McCarthy's contention that specialty goods somewhat resemble convenience goods from a shopping behavior point of view:

Shopping for a specialty good doesn't mean comparing but merely finding it. If such goods are readily available, their purchase may look like routine staple buying (McCarthy, 1975).

#### Conclusions and Implications

A rather unexpected finding in our research is that females evaluated the advertising in general higher than males. Several specific products ads benefited more from a female audience, even though no corresponding effect was found for comparison advertising versus non-comparison ads. Ex post facto this result certainly justifies our using a design which permitted isolating this source of variation, and it is obvious more in-depth research is needed to uncover why this phenomenon occurred. Perhaps there is a "congeniality" factor here that could emerge under more targeted research.

More importantly, our research findings would strongly question the current use of comparison advertisements by firms marketing convenience and specialty goods. These products do not require the explanation of differences between brands, quality, style, price, colors, features, and so on which is typical of comparison advertising messages. A firm, therefore, should examine what type of product it is marketing on the basis of the way consumers buy the product, before a comparison advertisement is selected instead of an individual product ad. If the firm is marketing a shopping good, a comparative ad represents a potentially powerful alternative.

For a large number of products, our findings do not support the FTC's claim that comparison ads provide consumers with information they desire before making a purchase decision. Instead, support was lent here to Ogilvy and Mather's conclusions concerning package goods advertisers. Still, our study, despite the strength of the research design and the random sample, is limited by the somewhat unsophisticated advertising medium and displays. Further research should investigate the same inter-relationships using more realistic media.

References

"Be Hard on Comparative Ads: Roberts to Four A's," Advertising Age, 44(November 19, 1973), 1.

M. Christopher, "NBC Widens Comparative Price Ad Rules," Advertising Age, 46(September 1, 1975), 42.

D.G. Clarke, "Sales-Advertising Cross-Elasticities and Advertising Competition," Journal of Marketing Research, 10(August, 1973), 250-261.

"Code Review Board Adds More Teeth to TV Rules for Comparison Ads," Broadcasting, 88(October 7, 1974), 61.

"Comparative Ads on Upswing, So Are Complaints," Broadcasting, 90(May 10, 1976), 51.

"Datril Ad Claims Based on Old Tylenol, Says J & J," Advertising Age, 47(April 26, 1976), 1.

P.H. Dougherty, "Calling Brand X by Its Real Name," The New York Times, (January 21, 1973), 15.

"Four A's Does an About-Face in Comparative Ad Guidelines," Television/Radio Age, (April 15, 1974), 49.

N. Giges, "Coca-Cola, Pepsi Put Taste Tests Into Comparison Arena," Advertising Age, 47(April 5, 1976), 1.

R. Giray, "Ford Extends Comparisons to Pinto Line," Advertising Age, 47(March 22, 1976), 4.

L.J. Hugh, "TWA Widens Comparative Ad Use Despite More Industry, CAB Pressures," Advertising Age, 47(March 15, 1976), 2.

L.J. Hugh, "TWA Again Revises On-time Ad, Starts Tackling Overseas Lines," Advertising Age, 47(March 22, 1976), 2.

R.E. Kirk, Experimental Design: Procedures for the Behavioral Sciences, Brooks/Cole Publishing Co., Belmont, California, 1969.

J.D. McCarthy, Basic Marketing: A Managerial Approach, 5th ed., Richard D. Irwin, Inc., Homewood, Illinois, 1975.

S. Misra, and S.C. Jain, "Effect of Fittingness, Type of Goods, and Type of Slogan on Brand Awareness," Journal of Applied Psychology, 35(November 6, 1971), 580-585.

J.L. Myers, Fundamentals of Experimental Design, Allyn & Bacon, Inc., Boston, 1969.

V.K. Prasad, "Communications Effectiveness of Comparative Advertising: A Laboratory Analysis," Journal of Marketing Research, 13(May, 1976), 128-137.

H. Scheffe, "A Method for Judging All Contrasts in the Analysis of Variance," Biometrika, 40(1953), 87-104.

"Schick, Inc., Teeters on the Razor's Edge," Business Week, (May 5, 1975), 38.

T.A. Shimp, "Comparison Advertising in National Television Commercials: A Content Analysis," Proceedings of the Fall conference of the American Marketing Association, 1975.

"So Long, Brand X - Naming Names of Rivals in Ads is Catching On But Spurs Controversy," The Wall Street Journal, 54(December 26, 1973), 1.

"Tannenbaum: Comparative Ads Can Work: Kershaw Says No," Advertising Age, 47(May 17, 1976), 1 and 110.

S.M. Ulanoff, "Comparison Advertising: An Historical Retrospective," Working Paper, Marketing Science Institute, 1975.

W.L. Wilkie, and P.W. Farris, "Comparison Advertising: Problems and Potential," Journal of Marketing, 39 (October, 1975), 7-15.

B.J. Winer, Statistical Principles in Experimental Design, 2nd ed., McGraw-Hill Book Co., New York, 1971.



## LEVELS-OF-PROCESSING, PERCEPTUAL BIAS, AND COMPARISON ADVERTISING

John G. Myers, University of California, Berkeley

This paper reviews three Frontiers in Advertising studies presented at the Annual ACR Conference, Miami Beach, Florida, October 26-29, 1978 by: Joel Saegert, "A Demonstration of Levels-of-Processing Theory in Memory for Advertisements"; Peter H. Webb, "Perceptual Discrepancies in the Time Decision and Number of Television Commercials"; and Edwin C. Hackleman and Subhash C. Jain, "An Experimental Analysis of Attitudes Toward Comparison and Non-Comparison Advertising." The review identifies common themes and basic similarities and differences, provides critical commentary on each, and concludes with a brief section on implications and directions for future research.

The three studies represent recent work on advertising effectiveness research from what might be called a consumer information processing perspective. Saegert's focus is on testing some aspects of a new theory of learning called "levels-of-processing." Webb's work is motivated by a general advertising problem referred to as "clutter," and Hackleman and Jain focus on a second advertising problem--the relative effectiveness of comparison versus non-comparison advertising. An interesting departure point is the nature of the criterion variable used in each case. In both the Saegert and Webb studies, cognitive or perceptual levels are of central interest. Saegert uses recognition and recall measures and Webb uses measures of commercial time and number of commercials recalled. Jain and Hackleman concentrate on affective or evaluative levels and use a criterion best described as preference for alternative types of advertising stimuli.

Another common theme is a basic similarity in methodological perspective. All three studies are causal/experimental research of some kind and test main and interaction effects of one or more treatment factors. Data collection is handled by manipulating stimulus materials (advertising copy) and/or the subject pool (e.g., sex), introducing various types of controls, and exposing subjects (consumers) to test materials in "laboratory" or simulated natural environment situations. Data analysis is largely confined to comparing group mean differences and analysis of variance. There is thus great methodological consistency across the studies even though the designs differ in important ways. None however is representative of other forms of consumer information processing research such as protocol research or tests of the predictive power of a preference function or theory (e.g., expectancy-value, weighted beliefs, conjoint analysis, constant sum paired comparisons). Direct rather than unobtrusive measures (e.g., response latency, chronometrics, facial action coding) are used to get at recall and preference in all of the studies.

The substantive questions addressed in each case differ, and there is an interesting difference in research perspective. The recent Commission on the Effectiveness of Research and Development for Market-

ing Management<sup>1</sup> proposed a classification in which all marketing research can be considered either: (1) Basic research; (2) Problem-oriented research; or (3) Problem-solving research. Saegert's study is probably close to what the Commission would call basic research, whereas Webb's and Hackleman and Jain's work appears more "problem-oriented." Some additional comments on this theme are given at the end of the review.

### A Demonstration of Levels-of-Processing Theory in Memory for Advertisements

The major contribution of this work is to test some aspects of "levels-of-processing" theory attributable to Craig and Lockhart and recognized for its potential in understanding processes which underlie advertising recall by Olson. The basic argument is that memory is a function of the "level" to which material is processed and is to some degree independent of the amount of repetition or rehearsal of the stimulus. As Saegert points out, a precise definition of "level" has not been specified, but in general it refers to the degree to which material is subjected to elaboration in relating it to a viewer's prior experience and knowledge. Although not discussed in the paper, the implicit assumption appears to involve a linear relation between the amount of elaboration and the amount of memory or recall. Readers should not have to reflect hard to appreciate the underlying theoretical controversy here between cognitivists who argue for the quality of processing as the major determinant of learning, and s-r behaviorists more prone to argue for the quantity of rewarding or nonrewarding stimuli and who emphasize over-time learning, trials, repetition, decay, and so on.

The operationalization of levels-of-processing is interesting and reveals much about the study. Thirty male and female adults were asked to answer questions about forty brands (from magazine advertisements). The "deep processing group" answered questions concerning their personal experiences with the brands such as, "Do you have this brand in your home?" The "shallow processing group" were questioned on formal aspects of the brand name: "Is the brand name in script letters?" Deep and shallow questions were reversed for the first and second half of the subjects. A recall measure (write down as many of the brand names as you can remember), followed by a recognition measure (identify test names from a list of 80 possibilities) were used as the two main criterion variables or factors. Analysis involved a 2 x 2 factorial with type of question as the repeated measures variable. In both recall and recognition, the differences between the deep and

<sup>1</sup>For a review article on the Commission's work, see John G. Myers, Stephen A. Greyser, and William F. Massy, "The Effectiveness of Marketing's 'R. & D.' for Marketing Management," Journal of Marketing, forthcoming, January 1979.

shallow question items were highly significant. There were no significant order or interaction effects between order and question type.

This is a nice straightforward piece of work in which an important new theory is tested. One can say that the null hypothesis of no significant effects of processing level is rejected at a highly significant level. We have here some new insights into that obtuse construct known as "involvement" in understanding advertising effects and learning.<sup>2</sup> The study stimulates a whole series of questions pertaining to the construct. Is the function linear or nonlinear? Is "personal elaboration" the same thing as Bogart's "connections"? Is Krugman's "low involvement" the same thing as "shallow processing"? Could deep processing lead to the wrong kind of learning (for example, if one equates deep processing with "counter-arguing")? How might distraction or refutational advertising play a role in understanding these memory effects? What is the impact of deep processing on attitude and attitude change or on subsequent brand choice? Theoretical and empirical work addressed to any of these questions would extend the Saegert study, and could provide useful new insights to the nature of the process. From the practitioner's viewpoint, the recommendation that creative people should encourage deep processing and personal elaboration is reasonable. It will not, however, come as new news to many creative directors!

Perceptual Discrepancies in the Time Decision and Number of Television Commercials

The term "clutter" has been used to refer to all the nonprogram elements on television including commercials, promotional announcements, public-service announcements, billboards, station breaks, and credits. It is a problem for advertisers because it potentially decreases the effectiveness of individual messages, for the broadcaster because it increases operating costs and absorbs program time, and for the viewer because frequent and lengthy program interruptions can generate frustration and interfere with program content.

Webb's paper reports on some special aspects of the clutter question, specifically whether or not consumers tend to exaggerate the amount of time actually devoted to commercials, and/or exaggerate the actual number of commercials aired in a particular time block. The paper is a spin-off of the clutter research project undertaken by Ray and Webb (1974, 1976)<sup>3</sup> sponsored by the Marketing Science Institute, and readers will find it useful to refer to the original references in understanding the broader issues involved. This article is mostly concerned with reporting the results of two questions that were asked during that project: How many commercials were there? (consider this the Number

<sup>2</sup>In an early formulation of a theory of advertising effects, personal involvement and stimulus relevance were considered the two principal processes associated with a perceptual filter conditioning the amount of attention or "psychological tension" aroused. See John G. Myers, Consumer Image and Attitude (Berkeley: Institute of Business and Economic Research, 1968), pp. 89-92.

<sup>3</sup>See Peter H. Webb and Michael L. Ray, "The Effects of Television Clutter: An experimental Investigation," Marketing Science Institute Project Description (1974); and Michael L. Ray and Peter H. Webb, "Experimental Research on the Effects of TV Clutter: Dealing with a Difficult Media Environment," Report No. 76-102, Marketing Science Institute, April 1976.

dimension); and How much time was devoted to commercials? (call this the Time dimension).

Motivation for this work stems from a general interest in learning more about the clutter problem (even though the authors report earlier that many advertisers consider it a "nonproblem"), and the more basic question of perceptual distortion generally. There is much evidence from surveys and public opinion polls that consumer beliefs are often at variance with "objective" facts, and Webb is interested in exploring this phenomenon in the case of perceptions of time and number of commercials. His literature review concerns evidence on whether the amount of time devoted to commercials and the actual number of commercials has actually increased (in general, the answer to both questions is yes, although the evidence is skimpy--standard commercial minutes in nonprime time, for example, increased from 9½ minutes in 1952 to 16 minutes in 1978), and evidence on possible determinants of perceptual distortion with respect to time. It is difficult to believe that experimental psychologists and others have not studied perceptual time distortion, as Webb implies, and that there is not some evidence on number distortion which is not covered in the review. There is, for example, a vast literature on "selective perception" in social psychology, and indeed the concept of "perceptual bias" is a part of most comprehensive models of buyer behavior. The germ of Webb's theoretical explanation is based on recent dissertations concerning stimulus complexity as the major explanatory variable. The more complex the stimulus, the greater the tendency to exaggerate the amount of time spent in studying it. He acknowledges that other factors such as prior expectations of the viewer, relevance of the stimulus, affect (pleasant/unpleasant stimuli) and novelty may contribute to an explanation of why distortions take place.

A brief review of the major results from the two studies reported in his paper is given below:

Study I	TIME	NUMBER
Baseline (zero comm.)	Minor Overestimate	Minor Overestimate
Low Clutter	Major Overestimate	Accurate
High Clutter	Minor Overestimate	Minor Underestimate
<u>Study II</u>		
By Program/Commercial Location (prime v. daytime)	Major Overestimate	Major Underestimate
By Commercial Schedule (6 x 2, 3 x 4, 2 x 6)	Major Overestimate	Major Underestimate

As can be seen, there appears to be a consistent pattern across the two studies in viewer tendencies to overestimate the amount of time devoted to commercials and to underestimate the actual number of commercials that were shown. What is more difficult to explain are the distortion patterns within and across the studies. In Study I, viewers more than doubled their estimate of actual time in the low clutter condition. In effect, the distortion was much greater in the low clutter than

in the high clutter condition, and this would appear not to support the complexity hypothesis. Also, the relative degree of accuracy in number estimation across the two studies is disturbing. In Study I viewers were much more likely to give accurate estimates, while in Study II these estimates were considerably below the actual number.

Webb's courage in reporting on two studies of similar things in the same paper is to be admired, and the cross-methods/cross-traits spirit which underlies it is encouraging. The paper above all raises interesting and pertinent questions associated with perceptual distortion and clutter, even though some results appear inconsistent and contradictory. The tendency to present hypotheses post hoc, to choose one type of analysis for one study (t-tests) and a different type for the other study (Anova), and to skirt the problems of operational pretesting of constructs like stimulus complexity should be avoided. Numerous competitive explanations for the results can be put forward. The accuracy of number estimation in Study I might be attributed to the greater ease of estimating number rather than time (but this is not well supported in the second study). Differences in prior expectations, experience, or amount of knowledge might explain some variance and suggests the need for covariance analysis. Sex differences in the samples (Study II was an all-female sample) may have been operating. The decision to keep as closely as possible to a simulated natural environment during the exposure situation (viewers could talk to each other, walk around, and so on) although normally a positive factor in raising external validity (where interest is focused on the effectiveness of the commercials), in this case may have led to other confounding effects (e.g., between those paying close attention and those merely guessing).

An Experimental Analysis of Attitudes  
Toward Comparison and Noncomparison Advertising

This study by Hackleman and Jain is interesting from the viewpoint of comparisons with the previous two because of its emphasis on experimental design and its comparative lack of attention to theoretical constructs, or theory development. For those interested in measurement and experimental design, this is an excellent exposition of the use of split-plot repeated measures factorials, latin squares, factor analysis, and other randomization techniques for testing and controlling various confounding factors. The authors have used exceptional care and painstaking effort in these aspects of their work. The principal limitations of the study relate to a lack of attention to theory and a tendency to over-generalize from the study results. Some generalizations do not in fact appear to be supported by the results presented in the paper.

The central question which motivates this work is that of the "effectiveness" of comparison advertising versus noncomparison advertising. The operational measure of effectiveness chosen can be called "preference for an advertisement." One can argue that this is not a very good measure if the viewpoint is that of the advertiser interested in building awareness, comprehension, or favorable attitude for his brand. Liking for an ad and favorable attitudes for the brand advertised are not the same thing. The other focus of the study is on possible determinants of relative effectiveness beyond advertising type, specifically sex and product, and the interactions between these three factors in predicting effectiveness. It might be said that the situation which motivates the study is that advertisers generally don't like comparison advertising, the FTC is encouraging it, and we know very little about what consumers like or want!

The authors' literature review is skimpy. There has been much more work done on comparison advertising than reported here. Also, the paper would be strengthened by recourse to some theoretical perspective to explain possible effects. What is a "comparison advertisement" from the viewpoint of cognitive, affective, or motivational components on the receiver side or source and message components on the sender side? Why should comparison advertising be more or less effective than noncomparison advertising in terms say of counterarguing, distraction, refutation, and so on? What would congruity, balance, dissonance (consistency) theories, levels-of-processing, high- or low-involvement theories, complexity theories, expectancy-value, attribution, or any other theory predict about likely effects?

The real value of this paper is in studying the operational procedures used in implementing the split-plot design. Great care was taken in developing stimulus materials. Free response word association techniques were used to develop new brand names for the 12 products tested. Comparison ads were developed by the clever device of combining two noncomparison ads, and locations of the two rotated in test materials. A 10-item preference scale was developed by factor analyzing a larger bank and choosing items that appeared to have highest reliability and construct validity (it is not clear how these choices were made nor why they resulted in higher reliability and validity). In addition to the 2 x 2 x 12 factorial with 120 replications (subjects), further controls were incorporated by stimulus rotation and latin square designs. The authors may be trying to "kill a fly with a bulldozer," but their experimental methodology is well worth careful study and review.

Some reported results unfortunately do not flow easily from the data. Consider, for example, the simplified recasting of their Table 5 reproduced below:

	Product Classification	Non-Comparison Message	Comparison Message
Coffee	Convenience	H	L
Headache Pill	Convenience	H	L
Cigarettes	Convenience	H	L
Camera	Shopping	H	L
Piano	Specialty	H	L
Vacation	Specialty	H	L
Liqueur	Specialty	H	L
Toothpaste	Convenience	L	H
Refrigerator	Shopping	L	H
Clock Radio	Shopping	L	H
Shaver	Shopping	L	H
Car	Shopping	L	H

H (high) and L (low) in the table refer to the direction of preference in each case. It is difficult to conclude as the authors did that comparison ads are more "effective" in the case of shopping goods, although the tendency (if effectiveness is assumed to be liking for the ad) is certainly there. There is, nevertheless, one case in which the shopping good non-comparison ad was preferred, and another case in which the convenience good comparison ad was preferred. Another stated conclusion is that there were no differences in attitude toward comparison and noncomparison ads. Although an F-test does support this overall, the

preceding table can be interpreted to mean that non-comparison ads are preferred in seven product cases and comparison ads preferred in five cases. The cleanest finding may be that sex appears to have an overall effect (women preferred all ads more than men), but did not discriminate preference for comparison and noncomparison ads. The implications or explanations of this phenomenon appear, however, not to be related to the principal motivations for the study.

Several other types of questions can be raised. Is the trichotomy of convenience, shopping, and specialty goods really useful here? Should not other criterion variables such as claim believability, comprehension, and so forth have been included given the nature of the controversy over comparison advertising? The preference measure chosen is only one of numerous possibilities, and probably not the best one. Summing items over a semantic scale assumes each item is equally important and this may not be the case. One or two other methods to get at preference and to test construct validity could well be included. Effectiveness is a nebulous concept in advertising. Much depends on the advertiser's objectives. Neither of the major conclusions and recommendations--that firms marketing convenience and specialty goods should avoid comparison advertising, and that there is no support for the FTC's claim that comparison advertising provides consumers with more information--are truly iron-clad generalizations which can be made from these results.

#### Implications and Directions

Possible extensions to each of the three studies were given implicitly or explicitly in the preceding sections. The most basic issue concerning research directions is the nature of the kind of research that needs to be encouraged. Marketing scholars too often appear faced with the dilemma of doing "great research on a trivial problem," versus doing "trivial research on a great problem" when the "problem" is specified in terms meaningful to the decision-maker or manager. Basic research is often equated with the first option and problem-oriented research with the second. Although the Commission study referred to earlier<sup>4</sup> concluded that all forms of research (basic, problem-oriented, and problem-solving) were vital and important to the development of the field of marketing, there was general agreement that the problem-oriented category is the kind that needs encouraging. It is not clear that this is good advice if the consequence is essentially a fostering of "trivial research" nor that a desired synthesis of "great research - great problem" perspectives is an attainable goal. There are the seeds here for some interesting debate on these issues.

---

<sup>4</sup>See footnote 1.

## Introduction

The three papers in this session illustrate quite dramatically the wide diversity of research in consumer behavior. All papers use laboratory experiments to investigate advertising; yet the papers address vastly different questions and approach their problems in quite different manners. The first paper by Joel Saegert attempts to apply a cognitive information processing concept to advertising recall. The second paper by Ed Hackleman and Subhash Jain, which is quite atheoretical in nature, contrasts consumer perceptions of comparison and non-comparison advertisements. The last paper by Peter Webb explores the factors which determine advertising "clutter" as measured by viewer perceptions of the amount of advertising in programming.

## A Demonstration of Levels-of-Processing in Advertising Recall

Saegert has presented one of the first empirical tests in a consumer behavior context of the "levels of processing" theory of Craik and his colleagues (Craik and Lockhart, 1972; Craik and Tulving 1975; Lockhart, Craik, and Jacoby, 1975). This theory, which focuses on the processes rather than the structures of learning and memory, has had a large impact on cognitive psychology and is very likely to lead to important research in consumer behavior. Saegert's research is important since it introduces many of us to this important area. (An excellent review of this area is available in Olson (1977)).

The goal of Saegert's research was to assess whether the results of experiments reviewed by Craik and Lockhart (1972) and/or conducted by Craik and Tulving (1975) could be replicated with advertisements instead of meaningful words as stimuli. Although his experiment seems to have been carefully conducted, I am a little uncertain whether he sufficiently replicated the procedures reported by Craik and his colleagues. An exact replication is not vital per se; however, I wonder exactly what was manipulated by Saegert's procedures. It may be that Saegert's hypotheses were confirmed but that the differences in recall and recognition were due to something other than his cited reason. At the very least, I worry that Saegert's experiment may lead to the misleading perception that the levels of processing notion is a rather simple one with according limited implications for consumer behavior research.

Saegert mentioned two similar paradigms of past research. The first by Hyde and Jenkins (1969, 1973) manipulated the type of orienting task to experimental groups. Some experimental groups were instructed to intentionally learn and remember the words, and others were asked to process the words in a certain way but were not told that there would be a subsequent learning test. Within the incidental learning condition, subjects were asked either to check whether the letter "E" appeared in the word, the number of letters in the word, or to rate the pleasantness of the words. The former two conditions were interpreted by Craik and Lockhart to be shallower than the latter. Another set of processing manipulations was reported by Craik and Tulving (1975). They instructed different groups of subjects to process at presumably increasingly deeper levels by either structure ("Is the word in capital letters?"), phonemic ("Does the word rhyme with WEIGH?"), categorical ("Is the word a type of fish?"), or sentence ("Would the word fit the sentence \_\_\_\_\_?") instructions.

The first question types were assumed to result in a structural processing, whereas the last two led to increasingly semantic processing.

In his experiment, Saegert asked subjects to process either structurally or semantically the brand name in a print advertisement. The structural manipulation very closely resembled previous ones. However, I am less certain of the result of the semantic questions. Saegert mentioned two examples: "Do you have this brand in your home?" and "Have you ever bought this brand?". Unlike Craik and Tulving's instructions (which were most similar to Saegert's), Saegert's questions were not able to anticipate whether the answer would be "yes" or "no." This may be an important aspect. Craik and Tulving found higher recall and recognition for "yes" than "no" answers in semantic processing modes such as judging whether a blank in a sentence could be filled by the target word. Also, more complex semantic processing instructions led to greater recall with questions with positive answers but not with negative ones.

This is important to the theoretical distinction between depth and spread or elaboration of processing. Craik and Tulving hypothesized that positive responses enable the subject to form a unified and more meaningful image of the complete cue. A congruent "yes" response may help to also recall other associations with the target stimulus.

Because this research is assessing the replicability of past cognitive psychology experiments, it might have been more appropriate to more exactly follow one particular procedure. For example, questions analogous to those used by Craik and Tulving might include "Is the advertised brand a physical product (as opposed to a service)?" or "Is the advertised brand something people eat?" Such questions would offer the advantage of being able to anticipate a "yes" or "no" and thus being able to control for congruity and coinciding greater elaboration in processing. Then the effects of positive and negative answers could be examined for similarities to past results. There may also be some theoretical importance to distinguishing between ego-oriented processing (e.g., "Is this brand in your home?") and more objective-processing. I'd expect that the former would lead to more elaborative processing by evoking more personal connections (Krugman, 1965). The greater recall and recognition in the deeper processing condition in Saegert's experiment may have been due to the personal nature of his questions and the resulting ego-oriented processing rather than or in addition to actual deeper processing. The two possibilities cannot be separated due to their confounding.

Saegert's "yes" and "no" answers could be compared for differences in subsequent brand name recall or recognition. Of course, with the subjective questions used by Saegert, any greater recall for positive answers could be alternatively interpreted as the result of greater initial familiarity. As Olson (1977) has emphasized, initial familiarity is a necessary condition for deeper, more semantic processing. For extensive semantic processing of a stimulus to occur, a subject must have a previously developed semantic structure of knowledge about that stimulus. Therefore, it was probably important that Saegert used adults as subjects who were more

apt than, say college students, to be familiar with the various advertising brands. It might be interesting to manipulate the type of processing of ads but to analyze the effects of processing within levels of familiarity or usage. One would expect that the effect of greater processing would be greater with brands initially familiar than those which are unfamiliar because the more elaborate reactions to deep processing instructions can be evoked only in the former condition.

Saegert's data may be analyzed in other ways to offer suggestions about future research. Certainly, as Saegert pointed out, it is very important for future research to ascertain whether and, if so, how advertising copy or formats can be designed to evoke deeper processing. It might be useful to include in an analysis of variance a term for ads which could test whether some brand names are better recalled than others. If so (which is likely), a post-hoc analysis of ad characteristics (e.g., long vs. short copy, questioning headline or not, etc.) in addition to brand familiarity might suggest how to manipulate recall. A similar procedure by Ray and Sawyer (1971) was somewhat productive in finding factors that interacted with repetition. Of course, recall differences may be due to factors other than induced processing but any insights at this early research stage would be welcome. I do not know if the five seconds per ad allowed by Saegert was sufficient for recall to be sensitive to different ad types. Certainly, the achieved recall levels (2 to 5 out of 40) were not very high.

Hopefully, processing can be manipulated by factors other than verbal instructions. Some positive evidence that this can be done was offered by Marslen-Wilson and Tyler (1976) who manipulated the level to which spoken prose passages could be processed by presenting either normal sentences, sentences which, although grammatically correct, were difficult to semantically process, or sentences with neither syntactic or semantic information. Recall of the sentences increased with the level of available processing. However, to be useful for applied purposes, more subtle manipulations within the verbal communication must prove effective.

A great deal of consumer behavior research has involved the process by which exposed information becomes transposed into attitudes and behavior. Levels of processing theory may suggest more efficient theoretical explanations of this process and point out profitable avenues for future research. For example, other than due to message characteristics, processing may vary on either a situational or individual basis. One difference that may be either individual or situational is involvement. The notion of consumer involvement has been much discussed (e.g., Krugman, 1965) but with very little theoretical base (see Houston and Rothschild, 1978). It seems obvious that Krugman's claim of low ad recall with low involvement can be explained by shallow or superficial processing.

Processing level might also be related to effects other than learning. Certainly the greater the extent of initial processing, the greater the opportunity for a consumer to compare new information with currently stored information. Cognitive responses, which are a possible measure of processing level, have been shown to be related to both immediate and delayed measures of attitude (Sawyer, in press). There is some evidence that negative cognitive thoughts are more persistent over time (Sawyer and Ward, 1977). Perhaps negative thoughts require more extensive processing and thus tend to be better retained. Caccioppo and Petty (in press) found that a message advocating a counter-attitudinal position elicited more thoughts relevant to the topic than a proattitudinal message. This result coincided with greater recall of the arguments in the

counter condition than in the pro condition. However, some method must be found to distinguish different intervening causes of greater recall. For example, greater retention of negative thoughts could be due to greater distinctiveness from most other thoughts instead of being due to deeper processing.

Certainly, as Saegert observes, repetition at least offers the opportunity of more elaborate processing (see Wyer, 1974). If repetition thus leads to greater learning, such learning may be related to affect change (see Stang, 1975). Grush (1976) has shown that the positive or negative nature of generated associations appears to mediate the direction of resulting affect. However, repeated meaningful communications may yield learning and affective reactions that are much more independent than for less meaningful stimuli such as nonsense syllables. There is evidence that, although there are similar patterns of learning and affect as the result of repetition, the two effects are not related to each other (Caccioppo and Petty, in press). Olson (1977) suggested a relationship in which learning and belief structure influence each other over time. Further research may help establish a more direct link between recall and attitudes by means of levels-of-processing.

#### Consumer Attitudes Toward Comparison Advertising

Hackleman and Jain report the results of an experiment about comparison advertising. As they point out, this topic is the subject of much concern by decision makers in advertising management and public policy.

The authors have obviously taken a tremendous amount of care in implementing their study. Ads were tested for twelve different products; subject subsamples from three different population centers were used; and extensive pretesting<sup>1</sup> on the fictitious brand names was conducted. The measurement of the ten rating scales randomized the order of the ten scales and whether the positive or negative pole of the adjectives was first. A complex experimental design was used which controlled for many factors (e.g., the balancing of brand names between the first and second position in the comparison ad copy). Finally, although I have trouble understanding how two separate ads could be combined into one comparative ad which in turn was similar to a third ad which was non-comparative, it seems likely that the ads were well constructed and quite comparable to real advertisements.

My major concern with this study involves its overall purpose. Hackleman and Jain state that the focus is on "effectiveness of comparison advertising from the viewpoint of consumers." Why is what consumers think about the advertisements very relevant? Although Leavitt's (1970) work has argued that various consumer perception factors may yield useful diagnostics, I know of no positive evidence that consumers' ad perceptions are a valid measure of effectiveness. Nor do I know of any advertisers or public policy makers who are primarily concerned with such measures. It seems to me that the key issues involve more "traditional" measures of effectiveness such as ad copy comprehension and recall, consumer use of comparative information in forming beliefs on compared product attributes, and brand preference.

In examining Hackleman and Jain's employed measures of consumer's attitudes toward the ads, I wonder what is actually being measured. It is admirable that the authors were explicitly concerned with reliability and construct validity, something most of us consider only when critiquing other people's research. However, I do not know how they assessed reliability and validity from the factor analysis of the pretest. Whatever the quality of individual scales, the combination of all ten

scales (presumably by summing) into one univariate scale makes no sense at all (e.g., Osgood, Tannenbaum, and Suci, 1957) and permits no detailed insights about the effects of comparison ads. I would advise a factor analysis of the ten scales followed by a multivariate analysis of variance of the resulting factor scores. Subsequent univariate analyses of variance could isolate which factors were significantly affected if the multivariate analysis indicates some significant overall effects.

Let me give an example of the potential usefulness of less ambiguous measures of ad perception. Wilkie and Farris (1975) hypothesized that comparison ads will be judged more informative and more interesting. Isolation of individual scales or factors that measure these dimensions could test Wilkie and Farris's hypothesis. Similarly, perceptions of confusion and believability might be of diagnostic value. However, as stated above, I question the relevance of even unambiguous measures of consumer attitudes towards ads.

It probably is a matter of personal taste, but I would prefer that subjects not be exposed to both comparative and non-comparative ads for the same product. Such a design may overly focus subjects' attention on differences between the two ads -- a big problem with the Ogilvy and Mather (1975) research. Arguments in favor of the within-subjects design include control of individual subject differences, as Hackleman and Jain mentioned, and more external validity if the eventual use of comparative ads for a given brand is combined with non-comparative ones (see Greenwald, 1976).

Finally, I thoroughly disagree with the conclusions of the authors about the effectiveness of comparative ads and about the differences by product type. Concerning the latter, I have always wondered about the operational distinction between shopping and specialty goods. How much inter-judge reliability is there in rating products into these two categories? Why are pianos or foreign liqueurs specialty goods and not shopping goods? Couldn't a camera be a specialty good? (Similarly, I cannot agree with the McCarthy quote about specialty goods; people do not merely explore to see if an overseas vacation or foreign sports car is available). Certainly Hackleman and Jain cannot be faulted for using a long used product classification scheme. However, I wish they had done a more formal statistical analysis on the product types. The authors included product type with twelve levels as one of three experimental factors in an analysis of variance. The significance of this factor means that there were differences among the twelve ad pairs. Only by analyzing type of "goods" (convenience, shopping, or specialty) and products nested within type of goods as separate factors could the claimed interaction of type of goods and ad copy (comparison or not) be isolated and statistically tested. The post-hoc analysis of Table 5 cannot be considered conclusive evidence. At best, it is suggestive.

Most important, Hackleman and Jain's conclusion that their "findings do not support the FTC's claim that comparison ads provide consumers with information they desire before making a purchase" has absolutely no foundation. Nor is there any evidence in support of Ogilvy and Mather's cited results about persuasion, brand identification, claim believability, awareness of competitors, or message confusion. How can any conclusions be made when most of the effects in the authors' conclusions were not even measured in this study? Inclusion of less ambiguous and more relevant measures that could have been the bases for such conclusions would have enabled the painstaking research of Hackleman and Jain to be much more useful.

There are many interesting questions to examine in future research about comparative advertising. An excellent research agenda has already been proposed by Wilkie and Farris (1975) who listed thirteen tactical issues and another thirteen hypotheses about the effects of comparison advertisements. I'll not repeat their ideas here. Rather than to discuss many additional specific research directions, I would like to suggest what I believe are appropriate types of research questions and a preferred process of generating research ideas.

I was impressed with two aspects of the Wilkie and Farris paper. First, there was an attempt to conceptualize likely sources of differences such as types of ads and products from various behavioral science and marketing theories. Second, they concentrated their thinking on more traditionally acceptable measures of advertising effectiveness such as ad attention and recall, correct brand identification, perceived brand position, cognitive responses, beliefs about brand's performance on product attributes, evaluation of attributes, and brand preference.

Prasad (1976) exemplified the type of research approach advocated by Wilkie and Farris. He looked at a tactical issue of managerial importance -- whether explicitly named competitor brands or whether the more traditional "brand X" should be included in the comparison ads. In addition, he tested a market factor of likely importance -- whether the advocated brand was most preferred by the audience or not. The tested hypotheses were generated from several concepts from social psychology such as indexing, selective learning and recall, source credibility, and attitude-discrepant communication. Measured communication effects included brand and claim recall (both immediate and one week delayed), claim believability, and perceived brand position. I mention Prasad's study not because of any special theories, designs, or measurement techniques but because it agreed with my bias toward Ray's (1978) advocated process of taking a managerially relevant problem, borrowing relevant concepts or micro-theoretical notions from behavioral or economic theory, and measuring appropriate communication effects.

Certainly other behavioral concepts such as attribution theory, refutational appeals, order effects of two-sided appeals, and assimilation-contrast theory could be used to generate other hypotheses. The value of using theory to generate hypotheses in comparison advertising research is that it helps to get away from the type of "Ad A versus Ad B" research common in applied research (see Ray, 1975) and thus may be able to make a contribution beyond the immediate question(s) about a particular type of comparison ad.

I am currently interested in the efficacy of a particular type of comparison advertising in which explicit comparisons are made on more than one product attribute. This strategy, suggested by Boyd, Ray and Strong (1972), attempts to reduce the determinance of an important or highly evaluated attribute by a comparison which suggests that all brands are equal in performance in that attribute. Such a strategy would allow a brand's superiority on another dimension to have a larger effect on brand preferences. Such research would also assess the wisdom of adding a third (not highly evaluated) attribute in which the advocated brand was not superior. While such research would test a specific comparative advertising tactic, it could also explore a more global communication problem of whether it is possible to change the contribution of an attribute to brand preference (e.g., Lutz, 1975).

## Perception of Advertising Clutter

Like Hackleman and Jain, Webb has investigated a problem of great interest to both advertisers and public policy makers. In his study, he has focused on a very interesting variable -- perception of time. Time perception and whether and how it may be manipulated should be of great interest to consumer behavior researchers in other areas as well as advertising.

There are some obvious and important strengths to Webb's studies. They certainly rate high in external validity. "Real" people were used as subjects; actual television commercials were used, and exposure was as natural as possible.

Webb's measures of perceived time length and number of commercials remind me of a similar venture of mine in my dissertation research (Sawyer, 1971). While studying the effects of repeated advertising exposures, I assumed that one effect of repetition is irritation and annoyance. I thought that a good measure of this effect might be the extent to which subjects overestimated the number of exposures as a function of the actual number and, perhaps, the type of ad that was repeated. As it turned out, the measure was subject to so much individual variance that it was not a sensitive measure of annoyance.

Webb has hypothesized that the degree of over- or under-estimation of length and number of commercials will vary as a function of their "clutter." My major concern with his paper involves my confusion about what is clutter. I've always thought of clutter as the number of ads per commercial break or the total number of ads in a program (Maneloveg, 1971). Webb agrees that the definition is not obvious and suggests an additional possible measure -- the number of commercial interruptions. Unfortunately, Webb's experiments have not helped me to decide which is the most appropriate definition.

Study one appears to define clutter as either the total number of ads or the number of ads per interruption. "Low" clutter was represented by two groups of two 30-second ads (four in total), and "high" clutter was represented by two groups of four ads (a total of eight ads). After being convinced that high clutter, not low, was the problem, I was somewhat surprised that low clutter resulted in greater over-estimation of the number and time duration of the ads.

Study two unconfounded the number of ads per interruption and total number of ads by controlling for the latter. Since I've always thought of increased clutter as the stringing together of more and more ads per interruption, I assumed that the experimental conditions of six sets of two 30-second ads, three sets of four ads, and two sets of six ads represented increasing levels of clutter. However, perhaps due to his uncertainty as to a proper definition of clutter, Webb did not choose to label the experimental conditions as high or low clutter. As in the first study, the second showed that, the lower the number of commercials per interruption, the greater the over-estimation of the total number and duration of the ads. Is less clutter "worse?"

In his description of the results of study two, Webb suggested that the number of total interruptions is a better measure of clutter, and that, given this definition, more clutter is worse.<sup>2</sup> To support the view that the number of interruptions is the key source of overestimating, an admirable attempt is made to conceptualize the underlying reasons for the over-estimates. Webb suggests that either complexity, differences from ad sequences that could be expected from past television viewing, or both could explain the

results of a direct relationship between the numbers of interruptions and the amount of over-estimation.

I find the complexity analogy difficult to accept. Why should six breaks of two ads be more complex than the other conditions? The rationale that "a string of unrelated non-program elements is likely to be judged more complex than an integrated program segment" doesn't make much sense since none of the conditions had a totally integrated segment. Since all elements -- ads and program -- appeared in each condition, the judgment of the complexity of the whole half hour sequence should not differ.

The expectation explanation seems more promising. If the number of 30-second commercials per break averages four in prime time, then the study one results would be consistent with an explanation that the greater over-estimation in the "low" clutter condition was due to subject's tendency to assume that approximately four and (not the actual two) ads were presented each break. Similarly, I'd expect that over-estimates in study two would be greatest where the number of ads per break was less than the expected number and lowest where the number of ads per break exceeded the expected. My rationale is based on my expectation that subjects can most easily remember the number of breaks, and, in calculating the total number of ads, subjects would multiply the number of breaks times the "normal" or expected number of ads per break. If my speculation is correct, this process should lead to differences in the estimates of the total number of ads. However, there are little differences in the number estimated and the direction of the slight differences shows the highest estimate in the 2 X 6 condition instead of the 6 X 2 condition.

Another way to have assessed the predictive validity of an expectation explanation would involve the manipulated program environment of either a prime time comedy show or a daytime serial. Webb states that the total number of commercials (twelve) is typical of daytime and exceeds the normal total for prime time. Therefore, due to departures from expected, overestimates should have been higher for the prime time than the daytime program environment. However, virtually no differences were found. However, if, as Webb seems to have decided, the number of commercial interruptions is the key factor in determining clutter, than the expectancy explanation would predict an interactive effect between program environment and commercial schedule. The interruption pattern of 6 X 2 is typical of daytime whereas three interruptions typify prime time. Thus, it would seem that over-estimates of the 6 X 2 schedule would be greater for the prime time environment than the daytime, whereas an opposite effect would be predicted for the 3 X 4 schedule. Since the 2 X 6 schedule is not representative of either environment, there should be no differences in the effects of that schedule between the two program types. Unfortunately, Webb did not analyze the interactive effects since, for some reason, he used one-way analyses of variance instead of a two-way analysis of variance.

I realize that this study of time perception was an exploratory one and was not the prime purpose of Webb's dissertation. Future research ought to try to resolve the question of the proper definition of clutter. I expect that one problem is that the definition of advertisers might differ from that of public policy people which might in turn differ from consumers. It seems to me that Webb has decided that clutter be defined from the latter perspective. However, if clutter is defined by whatever condition leads to highest estimates of the duration and number of commercials, any test of this relationship is tainted by the circular



definition.

Several issues are evident involving the measurement of time perception. Should it be the estimate of total time devoted to commercials, total number of commercials, or the ratio of one to the other? How do subjects arrive at these estimates? Do they, as I speculated earlier, do some form of mental arithmetic multiplying number of ads (actual or expected) by the number of breaks? Although it may not be easy to obtain valid results (e.g., Nisbett and Wilson, 1977), subjects could be asked to reconstruct how they arrived at their estimates. It seems reasonable that number of commercials is more easily perceived and remembered than their length. This was supported by Webb's results. Perhaps, estimates could be made more accurate if the questions broke down the memory task to the number of interruptions, the number of ads per interruption, and the length per ad.

Other measurement issues include questions about the construct validity of commercial time and/or number estimates as a measure of consumer annoyance or lack of enjoyment. Do perceptions of greater amounts of television time devoted to advertising lead to measurably greater irritation, and is such annoyance related in turn to less television viewing? Convergent validity could be assessed by examining the correlations between the time and number perceptions and measures of either the advertisements or total programming on such attributes as "complex," "annoying," "enjoyable," "too many commercials," "too many interruptions," "commercials are too long," "more commercials than usual," etc. Discriminant validity could be assessed by correlations with measures of ads and programming not expected to relate to perceived duration and number of commercials. Such additional measures would allow better tests of the theoretical explanations of the effects of ad schedules. The expectancy explanation might also be explored by examining individual differences. Webb tried to do this by measuring estimates of total program length but was thwarted by people apparently relying more on past experience than on actual perception. Perhaps analysis of only those who accurately estimated the program length -- whatever the reason -- would be fruitful. If these people were more likely to rely on expectation, then they might also be more apt to confirm hypotheses based on expectancy. Alternatively subjects might be grouped according to their television viewing habits or by their answers to questions, in a pretest unconnected to the experimental session, about their knowledge of break patterns and length of commercials per break.

It may be that the greatest value of this exploratory research is to raise issues about the underlying processes and how to manipulate time perceptions. Such a dependent variable has been usefully studied in other consumer behavior contexts such as time or distance to retail locations (e.g., McKay and Olshavsky, 1975). Further understanding of time perception may help in other areas. For example, negative attitudes about mass transit may be the result of misperceptions about the total time involved in comparison to driving in one's own car. Shoppers' hesitation to more carefully monitor product labels and prices may similarly be caused by over-estimates of the extra time necessary.

#### Conclusion

I enjoyed reading these three pieces of research about advertising, and I strongly commend the authors for their efforts. Many of my questions or criticisms may be due to my inability to understand from the brief descriptions exactly what was done rather than any actual deficiencies. I have tried to raise questions about the papers in hopes of isolating conceptual and

technical issues that can be addressed in future research. I am sure that the authors agree with me that any constructive discussion of the issues that they and I have raised will make their considerable research time and expense worthwhile.

#### Footnotes

- 1 I assume the authors included the product with the brand (i.e., Fantastic Camera). Otherwise, the pre-testing of the meaningfulness of the brand names would have made no sense to the subject. Also, the bi-polar adjective scale format of the meaningfulness scale more resembles a semantic differential than a Likert scale.
- 2 Although the perceived length per commercial was not calculated in study two, it showed a trend similar to the other measures with estimates of 71, 64, and 57 seconds per commercial in the 6 X 2, 3 X 4, and 2 X 6 conditions, respectively.

#### References

- H. W. Boyd, M. L. Ray, and E. C. Strong, "An Attitudinal Framework for Advertising Strategy," Journal of Marketing, 36 (1972), 27-33.
- J. T. Cacioppo and R. E. Petty, "The Effects of Message Repetition and Position on Cognitive Response, Recall and Persuasion," Journal of Personality and Social Psychology, in press.
- F. I. M. Craik and R. S. Lockhart, "Levels of Processing: A Framework for Memory Research," Journal of Verbal Learning and Verbal Behavior, 11 (1972), 671-84.
- F. I. M. Craik and E. Tulving, "Depth of Processing and the Retention of Words in Spisodic Memory," Journal of Experimental Psychology: General, 104 (1975), 268-94.
- A. G. Greenwald, "Within-Subject Designs: To Use or Not to Use?" Psychological Bulletin, 83 (1976), 314-20.
- J. E. Grush, "Attitude Formation and Mere Exposure Phenomena: A Nonartificial Explanation of Empirical Findings," Journal of Personality and Social Psychology, 33 (1976), 281-90.
- M. J. Houston and M. L. Rothschild, "Conceptual and Methodological Perspectives on Involvement," AMA Educators' Proceedings, 43 (1978), 184-7.
- T. S. Hyde and J. J. Jenkins, "Differential Effects of Incidental Tasks on the Organization of Recall on a List of Highly Associated Words," Journal of Experimental Psychology, 82 (1969), 472-81.
- T. S. Hyde and J. J. Jenkins, "Recall for Words as a Function of Semantic, Graphic, and Syntactic Orienting Tasks," Journal of Verbal Learning and Verbal Behavior, 12 (1973), 471-80.
- H. E. Krugman, "The Impact of Television Learning: Learning Without Involvement," Public Opinion Quarterly, 29 (1965), 349-56.
- C. Leavitt, "A Multidimensional Set of Rating Scales for Television Commercials," Journal of Applied Psychology, 54 (1970), 427-9.
- R. S. Lockhart, F. I. M. Craik, and L. Jacoby, "Depth of Processing, Recognition, and Recall," in J. Brown (ed.), Recall and Recognition, London: Wiley (1976).

- R. J. Lutz, "Changing Brand Attitudes Through Modification of Cognitive Structure," Journal of Consumer Research, 1 (1975), 49-59.
- H. D. Maneloveg, "Television Isn't Alone in Commercial Clutter -- Magazines, Newspapers Have It, Too," Advertising Age, October 11, 1971, 47.
- W. D. Marslen - Wilson and L. K. Tyler, "Memory and Levels of Processing in a Psycholinguistic Context," Journal of Experimental Psychology: Human Learning and Memory, 2 (1976), 112-9.
- D. McKay and R. W. Olshavsky, "Cognitive Maps of Retail Locations: An Investigation of Some Basic Issues," Journal of Consumer Research 2 (1975), 197-205.
- R. E. Nisbett and T. D. Wilson, "Telling More Than We Know: Verbal Reports About Our Mental Processes," Psychological Review, 84 (1977), 231-59.
- Ogilvy and Mather, Inc., "The Effects of Comparative Television Advertising that Names Competing Brands," 1975.
- J. C. Olson, "Theories of Information Encoding and Storage: Implications for Consumer Research," Working Paper No. 65, Pennsylvania State University, 1977.
- C. E. Osgood, G. J. Suci, and P. H. Tannenbaum, The Measurement of Meaning. Urbana: University of Illinois Press (1957).
- V. K. Prasad, "Communications Effectiveness of Comparative Advertising: A Laboratory Analysis," Journal of Marketing Research, 13 (1976), 128-37.
- M. L. Ray, "The Advertising Pretest as Part of a Multi-measure, Multimethod, Multisituation Validation and Application Research System" Advances in Consumer Research 2 (1975), 577-87.
- M. L. Ray, "The Present and Potential Linkages Between the Microtheoretical Nations of Behavior Science and the Problems of Advertising: A Proposal for a Research System," in A. J. Silk and H. L. Davis (ed.), Behavioral and Management Science in Marketing, New York: Wiley, (1978), 99-141.
- M. L. Ray and A. G. Sawyer, "A Laboratory Technique for Estimating the Repetition Function for Advertising Media Models," Journal of Marketing Research, 8 (1971), 20-9.
- A. G. Sawyer, "A Laboratory Experimental Investigation of the Effects of Advertising," Unpublished doctoral dissertation, Stanford University, 1971.
- A. G. Sawyer, "Repetition and Cognitive Response," in R. E. Petty, T. M. Ostrom, and T. Brock (ed.), Cognitive Responses to Persuasion, New York: McGraw-Hill, in press.
- A. G. Sawyer and S. Ward, "Carry-over Effects in Advertising Communication: Evidence and Hypotheses from Behavioral Science," in D. G. Clarke (ed.), Cumulative Advertising Effects: Sources and Implications, Cambridge, MA: Marketing Science Institute, (1977), 71-171.
- D. J. Stang, "Effects of 'Mere Exposure' on Learning and Affect," Journal of Personality and Social Psychology, 31 (1975), 7-12.
- W. L. Wilkie and P. W. Farris, "Comparison Advertising: Problems and Potential," Journal of Marketing, 39 (1975), 7-15.
- R. S. Wyer, Jr., Cognitive Organization and Change: An Information Processing Approach, Potomac, Md.: Lawrence Erlbaum (1974).

AN EXPLORATORY STUDY OF ASSERTIVENESS,  
AGGRESSIVENESS, AND CONSUMER COMPLAINING BEHAVIOR

Claes Fornell, Northwestern University  
Robert A. Westbrook, University of Arizona

Abstract

Authority, self-denial, personal guilt, and aggressiveness are embedded in traditional Western culture. This paper takes the position that these things inhibit effective communication and constructive criticism among members of society. The assertiveness model developed in Counseling Psychology is described and translated into a consumer context. The relationship between assertiveness/aggressiveness and consumer complaining is examined with some interesting implications for consumer policy.

Introduction

For the consumer, complaining is a means of making one's feelings known when unfair seller practices are encountered, when disappointment with a product arises, and when disapproval of business conduct more generally occurs. Yet the significance of consumer complaining behavior reaches considerably beyond these verbal expressions of dissatisfaction with the marketplace or its offering of goods and services. From an economic perspective, consumer complaints complement purchase choices as signals to producers to adjust the allocation of society's limited resources. From a managerial perspective, complaints represent potentially valuable information to guide marketing strategy. And from a public policy perspective, complaints may aid the development and targeting of consumer protection and market regulatory programs.

Despite the significance of consumer complaining behavior, it has only recently begun to receive attention in the literature. Empirical studies have focused mainly on determining the incidence of consumer complaining and its variation across demographic groups (Warland et al., 1975; Liefeld et al., 1975; Best and Andreasen, 1977; Pfaff and Blivice, 1977). There have also been attempts to relate complaining behavior to consumer perceptions of unfair selling practices (Kraft, 1977, Zaltman et al., 1977). While progress has been made in identifying selected relationships, it is evident that much remains to be learned about the determinants of consumer complaining.

Complaint behavior is often thought to be related to personality (Landon, 1977), but empirical studies have rarely addressed this issue. Moreover, available evidence is inconclusive. Zaichowsky and Liefeld (1977), for example, attempted to differentiate consumers who had written letters of complaint to a government agency from non-complainers using Cattell's Sixteen Personality Factors battery. Only very modest relationships were found, and the authors concluded that complaint letter writers were not acting primarily on the basis of inherent personality type was prone to complaining. In contrast, Wall, Dickey, and Talarzyk (1977) found that propensity to complain was related to a personality-like factor extracted from a set of ad-hoc AIO statements.

In neither of the preceding studies were the expected relationships between personality constructs and complaining explicitly conceptualized and tested; rather, relationships were sought from large general personality inventories. Jacoby (1969) and Kassarian (1970) have demonstrated the deficiencies and pitfalls in such an approach. Jacoby's criticism from a decade ago is equally valid today regarding the research on personality and consumer complaining. His point was that most investigators

operated without theory and with no a priori thought as to how and why personality should be related to the aspect of behavior under study. Not only did this result in little understanding, according to Jacoby, but as he showed by reanalyzing a classic study of the personality differences between Ford and Chevrolet owners, there was also a risk of drawing erroneous conclusions.

In one of the few other studies presenting evidence on the relationship of consumer complaining to personality, Faricy and Mazis (1973) developed a measure of complaint tendency as part of the overall construct of consumer dissatisfaction and found it related to Rokeach's dogmatism scale. They failed to find complaint tendency related to locus of control, however. This latter finding was supported by Zikmund and Miller (1974); though their measure of complaining was participation in a retail boycott.

Further study of consumer complaining behavior and its relationship to personality constructs is clearly warranted. In contrast to previous studies, greater theoretical conceptualization should precede attempts to relate personality measures and complaining. Moreover, the explanatory potential of other promising personality constructs should be examined. This paper describes such an exploratory study of the relationship between complaint tendency and the personality traits of assertiveness and aggressiveness.

Theory

Responses to Dissatisfaction and Frustration

The study of how people react to problems that arise when their ways of achieving satisfaction is prevented or interrupted is a major field of Personality Psychology. It is sometimes claimed that this is indeed the essence of personality, since it deals with the ways in which people cope with the barriers of goal-directed behavior and the frustration that result when expectancy is violated.

Frustration occurs when goal-directed behavior is blocked or interrupted before its completion (Mischel, 1971). Inasmuch as consumer behavior is goal-directed, frustration develops when the goal actually attained is something less than the goal sought, or when the goal sought requires more resources (money, time, energy) than the consumer is able, willing, or expecting to spend in order to achieve the goal. Frustration will also occur when the means of achieving satisfaction, including the resources as well as the goal-object (e.g., a product or a brand), are reduced or removed. Consequently, frustration can develop in both post-purchase and pre-purchase situations.

One of the earliest conceptualizations of frustration reactions was the frustration-aggression hypothesis (Dollard et al., 1939) stating that aggressive behavior not only assumes frustration, but also increases in probability with the occurrence of frustration. More recent developments, most notably in Learning Theory (Bandura and Walters, 1963; Berkowitz, 1969), have proven this hypothesis to be of questionable validity. Aggression is not seen as an inherent or "natural" reaction to frustration according to these theorists. Instead, aggression is viewed as learned response patterns. Furthermore, in the past few years there has been an increasing

interest in the concept of assertion and the distinction between assertion and aggression (Galassi and Galassi, 1977; Alberti and Emmons, 1974; Jukobowski, 1973). Concomitantly, there has been a growing interest in assertive training among the general public and, perhaps most notable, among women's groups. The purpose of such training is to help people communicate more effectively their feelings and opinions in all kinds of social and professional settings. There is also a rapidly expanding literature on the effects of assertive training.

According to Wolpe (1969), assertive training will help persons who have unadaptive responses. The training programs essentially include the following steps: (1) situation appraisal to determine what the rights and responsibilities are of the various parties involved and the probable consequences of various courses of action; (2) experimentation with new behaviors and attitudes in practice situations; (3) behavior evaluation to determine experienced anxiety, verbal content and delivery of message, and overall performance; and (4) behavior implementation (Galassi and Galassi, 1977).

#### The Assertiveness Model

Assertive behavior has been defined as "that complex of behaviors emitted by a person in an interpersonal context which express that person's feelings, attitudes, wishes, opinions or rights directly, firmly and honestly while respecting the feelings, attitudes, wishes, opinions, and rights of other persons." (Galassi and Galassi, 1977, p. 233.) According to these writers, assertion does not involve an undue or excessive amount of anxiety or fear. It represents the standing up for one's legitimate rights without violating the rights of others.

Mental health researchers of today appear to be in accord that assertive behavior is (a) learned and (b) situationally specific (Alberti and Emmons, 1974; Galassi and Galassi, 1977; Hersen et al., 1973; Jukobowski, 1973). As perhaps first suggested by Andrew Salter in his book Conditioned Reflex Therapy (1949) and now empirically demonstrated (Hersen et al., 1973), assertive training has the capability to reciprocally inhibit anxiety. Also, it is clear that assertiveness depends on the situation as perceived by the individual. The situational non-asserter may be cognizant of the appropriate course of action, but for one reason or another may choose to ignore it.

The person who impels his desire for self-assertion to excessive proportions by expressing his opinions in a hostile, threatening, or assaultive manner is aggressive. The aggressive person shows little or no consideration for the rights of others. The person who behaves aggressively does not recognize the potential consequences of his action and does not assume responsibility for them. It is for these reasons that aggressive behavior often results in unfavorable consequences for the aggressor as well as for the object of aggression. By contrast, assertive behavior is expressed with consideration of mutual rights and the possible outcome that may follow. Accordingly, the assertive person has a better chance of obtaining satisfactory remedy in a situation where expectancy has been violated.

Research on the effects of assertive training has reported increased self-esteem, increased positive reaction from others, and reduced anxiety in social situations for persons having gone through a training program. However, most research has dealt with treatment for sexual deviations (Stevenson and Wolpe, 1970; Edwards, 1972; Lazarus, 1971), marital problems (Fensterheim, 1972), socially anxious college students (Hedquist and Weinhold, 1970), and chronic schizophrenics (Weinman et al., 1972), which are areas that may be difficult to generalize to a consumer context. Although Alberti and Emmons in their highly influential work (1974) describe some consumer

situations among what they refer to as typical situations in which assertive behavior is called for, there has to our knowledge only been one published research study on assertiveness in a buyer-seller context. McFall and Marston (1970) assessed the effects of an assertive training program by having subjects treated experimentally and exposed to telephone calls where a "salesman" (experimenter) used several "high pressure" sales tactics to persuade subjects to subscribe to magazines. The differences between control and experimental groups were found to be relatively weak but in the expected direction.

Making the transition to a consumer dissatisfaction context, drawing upon the model of assertive behavior developed in Psychology and Mental Health and generalizing from the positive results of assertive training, what can be said about consumer complaining behavior? For one thing, it would be expected that the dissatisfaction experienced by a non-assertive consumer will produce undue anxiety so that the range of available remedy options becomes restricted. Since brand shifting, purchase or patronage termination, or total inaction are responses that do not require intercommunication, one would expect that complaining, which is an intercommunication initiated by the consumer, to be less likely a response of the non-assertive consumer. This is in contrast to the assertive individual who would not let possible unpleasantness prevent him from complaining to a company, if he thought that complaining was the appropriate course of action. According to the assertiveness model and the definitions of assertiveness/aggressiveness, the reactions to frustration would be based on different considerations for different individuals:

Response of an assertive individual =  $f(\text{available alternatives and probable outcomes})$

Response of a non-assertive individual =  $f(\text{the amount of anxiety associated with the alternatives})$

Response of an aggressive individual =  $f(\text{the magnitude and intensity of frustration})$

Only in assertive behavior are the consequences of the action fully considered beforehand. The aggressive person is more apt to consider the consequences after the action is taken, and the non-assertive person's main concern is to find a response with a minimal amount of intercommunication and anxiety. Consequently, one would expect that both assertive and aggressive individuals would be more likely than non-assertive individuals to complain to a company if dissatisfied with the product or service. If this is correct, it suggests some interesting implications for consumer policy. The purpose of this study was, then, to examine the relationships between assertiveness, aggressiveness, and consumer complaining.

#### Method

One hundred nineteen undergraduate students at a major private university participated in the research study by completing a self-administered questionnaire. Respondents were chosen selectively to represent all class levels and both sexes. The exploratory nature of the study, convenience, and economy were important but not the only considerations in determining the restriction of the sample to college students. The homogenous nature of the student population at this university substantially reduces the variability of potential exogenous influences beyond the scope of study. Moreover, the focus was on the relationships between relative scores rather than their absolute levels as descriptive of a total consumer population. It is difficult to find a priori reasons why the relationship between assertiveness, aggressiveness,

and complaining behavior would be any different between students and the general population.

Except from the benefits of having a pre-specified model guide analysis and selection of measurement items, the methodology used follows the pattern of the numerous studies on life-style and psychographics, utilizing factor analysis and regression. Respondents were administered the following instruments: (1) a set of measures concerning consumer complaint behavior and (2) a battery of items to assess individual assertiveness/aggressiveness. The differences in the possible sources of consumer dissatisfaction and targets of complaint communications suggested that several dependent measures of complaining be used. In this study, the following different complaint behaviors were examined:

1. Complaining to the manufacturer when the product fails to meet prior expectations.
2. Complaining to the retail store when a product fails to meet prior expectations.
3. Complaining to the manufacturer when a favorite product is discontinued.
4. Complaining to the manufacturer when a favorite product deteriorates in quality.

In order to reduce the influence of differences in product type, all of the preceding situations were limited to grocery-store items. Respondents indicated the extent of their agreement or disagreement on a six-point scale with each of the above situations.

While there is consensus in the literature on the conceptualization of assertiveness and its acquisition through learning, there is no similar agreement on measurement. Self-reports, behavioral responses, as well as physiological measures have been used with varying degrees of success (Hersen et al., 1973). Little has been done in terms of formal validation. In the present study, subjects were asked to indicate to what extent they agreed or disagreed with nineteen selected statements of assertiveness/aggressiveness, again on a six-point scale. The statements were taken from Alberti and Emmons (1974) and Evans (1977). Following factor analysis of the items, the factor scores were used as explanatory variables in regression analysis.

### Findings

All nineteen assertiveness/aggressiveness variables were submitted to a principal components analysis. Inspection of the distributions for each individual item did not reveal any abnormally skewed variables. A minimum eigenvalue criterion of 1.0 led to a seven-factor solution accounting for 61.9% of the variation in the original variables.

Although assertive behavior is clearly distinct from aggressive behavior according to the model presented, it does not follow that the two behaviors represent completely independent personality traits. We are not talking about dichotomous constructs -- it is possible to be assertive with some amount of aggressiveness included. In the final analysis, it is up to the subjective judgment of each party involved to determine whether or not a specific act is to be considered assertive or aggressive. Hence, the purpose of the principal components analysis was to minimize the cross-products of the factor loadings on the reference axes in order to obtain simpler factor structures without the restriction of orthogonality. Therefore, the factors were rotated to a fairly oblique terminal solution. The results are summarized in Table 1. Setting the

obliqueness parameter (Harman, 1967) equal to zero yielded relatively low factor correlations: over half were less than .10 and only one reached an  $r$  of .20.

Examination of the factor pattern in Table 1 suggests that factor I represents a dimension of "submissiveness," a definite form of non-assertiveness. Interpreted with consideration to directionality, the most cogent variables are (1) avoidance of people for fear of embarrassment, (2) finding it hard to say "no" to salesmen, and (3) having difficulty maintaining eye contact in a conversation. Factor II appears to capture "vociferousness" or aggressive self-assertion: a tendency to assert oneself in an aggressive or belligerent manner. This interpretation is suggested by the following high weights: (4) being openly critical of others; (5) calling unfair behavior to the attention of the offender; (6) showing anger by name-calling or obscenities; and (7) speaking out in protest when someone takes one's place in line. Factor III includes what appears to be "congeniality" or "high regard for human beings" but is difficult to interpret when the whole set of weights is considered. The three variables most closely associated with factor III are (8) having no difficulty complimenting or praising others, (9) finding no difficulty maintaining eye contact in a conversation, and (10) disagreeing with the statement that bullfight watchers should be given a taste of the suffering of the bull. However, (13) feeling at times so angry that one could resort to physical assault is also represented by the factor. Inspection of the structure matrix of correlations (not reproduced here) yielded similar results and did not provide any additional guidance. While this factor represents a mixture of variables that complicates its interpretation, on the whole, it does not include much assertiveness nor aggressiveness. Factor IV includes aggression with undertones of violence, based on fear rather than assertiveness. Its associated variables are: (11) believing that there should be a gun in every home; (12) viewing man as a dangerous and aggressive animal which is slowly becoming civilized; and (13) feeling at times so angry that one could resort to physical assault. Factor V displays a high weight (.88) in connection with only one item, (14) favoring strict enforcement of all laws. In considering the total structure as well as the pattern of this factor, it seems to lean slightly toward the non-assertive side of the continuum. Factor VI is closely associated with two items, (15) believing that slow drivers are not a greater menace than fast drivers, and (16) feeling that the U.S. would not be better off without "freaks." This pattern is suggestive of a dimension of tolerance versus intolerance. Factor VII finally appears to represent extent of "shyness," as distinct from "submissiveness" (factor I). The relevant variables of this factor are (17) being reluctant to speak up in a discussion, (18) not being the first to start a conversation with a stranger, and (19) feeling uncomfortable expressing opinions to an authority figure.

Table 2 presents classification of the factors according to their portrayal of aggressiveness and assertiveness. Some of the factors (F2, F7) include high or low levels of both traits, while one factor (F5) does not seem to represent either one. Using the assertiveness model as a criterion of construct validity, F1 and F4 come out ahead. Although the assertiveness model does not postulate completely dichotomous traits, it clearly distinguishes between assertive and aggressive behavior. F1 and F4 are the only factors that include either aggressiveness or assertiveness, but not both.

The results of the regressions, using the factors as explanatory variables, are found in Table 3.  $R^2$ 's varied from .08 to .16, a reasonable result in view of the numerous studies on personality and behavior reporting similar or lower ratios of explained variance. Just as in other aspects of consumer behavior, there are of

TABLE 1  
 FACTOR PATTERN MATRIX AND COMMUNALITIES OF ASSERTIVENESS/AGGRESSIVENESS ITEMS:  
 PRINCIPAL COMPONENTS WITH OBLIQUE ROTATION

Items	Factors							Communalities a)
	I	II	III	IV	V	VI	VII	
I often avoid people or situations for fear of embarrassment (1)	.67	.22	-.03	-.06	.15	.15	.28	.63
When a salesman makes an effort, I find it hard to say no (2)	.56	-.14	.11	-.05	-.08	-.29	.13	.52
I find no difficulty in maintaining eye contact, keeping my head upright in a personal conversation (3) (9)	-.60	.12	.41	-.08	.11	.02	.10	.57
I am openly critical of others' ideas, opinions, behavior (4)	-.16	.71	-.17	.03	.09	-.20	-.23	.68
When a person is highly unfair, I call it to his/her attention (5)	-.25	.65	.06	-.15	.08	-.12	-.10	.55
I sometimes show my anger by name-calling or obscenities (6)	.17	.61	.15	-.02	-.14	.00	.04	.44
I speak out in protest when someone takes my place in line (7)	-.30	.58	-.27	.08	-.02	.32	.06	.64
I find it difficult to compliment or praise others (8)	.30	-.08	-.72	.00	.00	.03	-.03	.61
People who watch bullfights ought to be given a taste of the suffering the bull has to experience (10)	-.22	.26	-.65	.04	-.06	-.21	.22	.64
There should be a gun in every home (11)	-.04	-.24	-.04	.80	.22	-.04	.12	.74
Man is a dangerous and aggressive animal who is slowly becoming civilized (12)	-.05	.05	-.06	.69	.33	.06	-.12	.62
Sometimes I can feel so angry or annoyed at a person that I feel I could hit him/her (13)	.17	.33	.39	.42	-.03	-.09	.09	.52
I am in favor of very strict enforcement of all laws (14)	.00	.06	.05	.08	.88	.16	-.08	.77
Slow drivers are more of a menace on the roads than fast drivers (15)	-.05	.16	.07	.03	-.16	-.83	.02	.72
The U.S. would be better off if there were no freaks (16)	.13	.06	-.03	.35	.33	-.52	-.13	.63
I am reluctant to speak up in a discussion or debate (17)	.11	-.03	.14	.05	-.04	-.29	.71	.70
When I meet a stranger, I am usually the first to begin a conversation (18)	.04	.08	.06	.00	.07	-.15	-.74	.57
I feel uncomfortable stating my views to an authority figure (19)	.35	.02	-.23	.03	.18	.00	.55	.64
I dislike arguing with people	-.12	-.32	-.04	-.25	.34	-.36	.26	.59

a) Before oblique rotation

TABLE 2  
 FACTOR CLASSIFICATION OF AGGRESSIVENESS/ASSERTIVENESS

Factor	Aggressiveness	Assertiveness
F1	-	Low
F2	High	High
F3	Low	Medium
F4	High	-
F5	-	-
F6	Low	Fairly High
F7	Low	Low

course many variables other than certain personality traits that affect complaining behavior.

Factor 1, which reflects "submissiveness," is also one of the factors in which more confidence of validity is placed (Table 2). It is significantly related to all complaint situations. Factor 4 (also with a high validity score), which was interpreted as "aggressive self-assertion," is significant in two of the regressions; all, however, show a positive relationship. Among other explanatory variables, it is worth noting that factor 6, which seems to capture a "tolerance" dimension and also scores relatively high on the validity criterion, is represented in two of the regressions. Factor 5, which

TABLE 3

## COMPLAINING MEASURES REGRESSED ON FACTOR SCORES

	Beta Coefficients			
	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>
F1	-.21*	-.21*	-.27*	-.19*
F2	-.10	-.02	.17*	-.05
F3	-.15	.03	-.07	-.13
F4	.17*	.10	.07	.29*
F5	-.08	-.13	.00	-.03
F6	.16*	.12	.06	.19*
F7	.08	.03	.10	.05
	R <sup>2</sup> =.12*	R <sup>2</sup> =.08	R <sup>2</sup> =.12 <sup>a</sup>	R <sup>2</sup> =.16**

a) .10 level significance

\*) .05 level of significance

\*\*) .01 level of significance

- 1) Complaining to manufacturer when expectations are unmet
- 2) Complaining to store when expectations are unmet
- 3) Complaining to manufacturer when favorite product is discontinued
- 4) Complaining to manufacturer when favorite product has deteriorated

does not include assertiveness nor aggressiveness, and thus has a low validity, does not contribute much to the variance across the complaint measures.

## Discussion

The regression results indicated that the factor labeled "submissiveness" was the strongest inhibitor of complaining. This finding was consistent across all situations examined. The a priori notion that the non-assertive consumer is less likely to resort to a complaint action as a means of alleviating frustration is thus supported by the data. Some other minor relationships suggested by the regression equations are more difficult to interpret, however. The fact that as many as seven factors, some of which are hard to label, emerged from a principal components solution only reiterates the fact that the assertiveness construct does not yet have the benefit of a well-developed measurement methodology. In the absence of rigorous validity assessment, which was beyond the scope of this investigation, the results are tentative and should be regarded as exploratory. In particular, this is the case for the minor facets of the relationships whose interpretations do not relate to the assertiveness model. For example, why is complaining to a manufacturer when a favorite product is discontinued "better explained" by aggressiveness and assertiveness than are other forms of complaining? Both aggressiveness and assertiveness account for more of the variation in situations of satisfaction blockage due to product unavailability, than they do when frustration is caused by product failure in meeting expectations. Contrary to what might have been expected, complaining to retail store displayed the weakest relationship to the assertiveness/aggressiveness variables. In view of the uncertainty regarding the amount of variance due to the instrument of measurement, it seems best to refrain from interpreting these relationships. To do otherwise would involve excessive conjecture.

Further inquiry into the interrelationships between the ways in which consumers struggle with frustration and the personality traits of assertiveness and aggressiveness is called for. If we want to learn more about the grass-

root dimensions of consumerism, the reasons why and the conditions under which some consumers voice their dissatisfaction, while others passively seem to accept or adjust to deficiency or injustice, we have to go beyond such surface variables as socioeconomic characteristics, political commitment, and consumer awareness of unfair practices and of channels for redress. More attention has to be given to the covert mental processes and the more or less enduring personality traits that affect frustration response. Even though consumer dissatisfaction is a case of interrupted or obstructed goal achievement, and several studies have examined how consumers deal with it, only fragmented attempts have been made to establish the link between frustration response and personality in consumer research. While the results of this first attempt to relate assertiveness to consumer complaining behavior posit a relationship that is static, direct, and linear, it seems likely that future studies may better depict the relationships involved using more complex arrays of associations. In the context of consumer behavior, assertiveness may be conceived of as an intervening variable whose function is to moderate the effects of other influences such as level of dissatisfaction, importance of problem, and awareness of remedy channels (cf. Nakaniski, 1972), and which probably interacts with these variables as well.

## Implications

Since assertiveness constitutes a learned behavior (as demonstrated in the psychiatric and psychological literature), and if non-assertive individuals consciously limit their remedy options by avoiding complaining (as suggested by the assertiveness model and the empirical results reported in this study), it follows that consumers can be taught through assertive training to better handle conflicts with sellers and manufacturers. This applies to the aggressive as well as non-assertive consumers.

It is well known that consumer policy cannot solely rely on consumer information (i.e., factual information about brand and product characteristics). It has to be complemented with consumer education (i.e., instruction on where to get and how to use available information) in order to have some impact. When both these methods fail to accomplish the desired goals, there are often calls for more consumer protection or market regulation. The assertiveness model suggests yet another way of improving the consumer's position in the marketplace. Its implication is that even well-informed, educated consumers may fail to effectively articulate their grievances because of undue aggressiveness or non-assertiveness. In such cases, it is clear that consumer information and education policies do not suffice, and protection and regulation may miss the mark. Since most consumer dissatisfactions are not voiced, and research has shown that about one in every five purchases results in some form of consumer problem (Best and Andreasen, 1977), assertive training for consumers may be a viable addition or alternative to current consumer programs. Assertive training for consumers does not mean expensive psychological counseling. Although there is, to our knowledge, no programs that are specifically tailored to consumers, many universities now offer courses in assertive training in their extension programs. There are also "do-it-yourself" manuals available.

Authority, self-denial, personal guilt, and aggressiveness are parts of traditional Western culture. They constitute some of the building material of our social structures and relationships, including the institutions and transactions in the marketplace. The assertiveness model highlights some of the negative consequences caused by too close an adherence to such values and suggests a way for emancipation. To further refine the role of assertiveness in a consumer context should be an interesting task.

## References

- R.E. Alberti and M.L. Emmons, Your Perfect Right: A Guide to Assertive Behavior, 2nd ed. (San Luis Obispo, Ca.: Impact, 1974)
- A. Bandura and R. Walters, Social Learning and Personality Development (New York: Holt, Rinehart and Winston, Inc., 1963)
- L. Berkowitz, "Control of Aggression," in B.M. Caldwell and H. Ricciuti, eds., Review of Child Development Research, 3 (1969).
- A. Best and A. Andreasen, "Consumer Response to Unsatisfactory Purchases: A Survey of Perceiving Defects, Voicing Complaints, and Obtaining Redress," Law and Society, 11 (Spring 1977), 701-742.
- J. Dollard, D.W. Doob, N.E. Miller, O.H. Mowrer, and R.R. Sears, Frustration and Aggression (New Haven: Yale University Press, 1939)
- N.B. Edwards, "Case Conference: Assertive Training in Case of Homosexual Pedophilia," Journal of Behavior Therapy and Experimental Psychiatry, 3 (1972), 55-63.
- C. Evans, Understanding Yourself (New York: A&W Visual Library, 1977)
- J.H. Faricy and M.B. Mazis, "Personality and Consumer Dissatisfaction: A Multi-dimensional Approach," in T. Greer, ed., Proceedings of the 1974 American Marketing Association Conference (Chicago: AMA, 1974), 202-205.
- H. Fensterheim, "Assertive Methods and Marital Problems," in R.D. Rubin et al., eds., Advances in Behavior Therapy (New York: Academic Press, 1972), 13-18.
- M.D. Galassi and J.P. Galassi, Assert Yourself -- How to be Your Own Person (New York: Human Sciences Press, 1977)
- H.H. Harman, Modern Factor Analysis (Chicago: The University of Chicago Press, 1967)
- F.J. Hedquist and B.K. Weinhold, "Behavioral Group Counseling with Socially Anxious and Unassertive College Students," Journal of Counseling Psychology, 17:3 (1970), 237-242.
- M. Hersen, R.M. Eisler, and P.M. Miller, "Development of Assertive Responses: Clinical Measurement and Research Considerations," Behavior Research and Therapy, 11 (1973), 505-521.
- J.J. Jacoby, "Towards Defining Consumer Psychology: One Psychologist's Views," paper presented at the American Psychological Association, 77th annual convention, Washington, D.C., September 1969.
- S.P. Jukobowski, "Facilitating the Growth of Women Through Assertive Training," The Counseling Psychologist, 4:1 (1973), 75-86.
- H.H. Kassarian, "Personality and Consumer Behavior: A Review," Journal of Marketing Research, 8 (November 1971), 409-414.
- F.B. Kraft, "Consumer Perceptions of Business Fraud: A Key Variable in Complaint and Repatronage Behavior," paper presented at the Southwestern Marketing Association meeting, New Orleans, March 23-26, 1977.
- E.L. Landon, "A Model of Consumer Complaint Behavior," paper presented at the Research Symposium on Consumer Satisfaction/Dissatisfaction and Complaining Behavior, Indiana University, Bloomington, April 21-22, 1977.
- A.A. Lazarus, Behavior Therapy and Beyond (New York: McGraw-Hill, 1971)
- J.T. Liefeld, F.H.C. Edgecombe, and L. Wolfe, "Demographic Characteristics of Canadian Consumer Complainers," Journal of Consumer Affairs, 9 (1975), 73-80.
- W.D. Martinson and J.P. Zerface, "Comparison of Individual Counseling and a Social Program with Nondaters," Journal of Counseling Psychology, 17:1 (1970), 36-40.
- R.M. McFall and A.R. Marston, "An Experimental Investigation of Behavior Rehearsal in Assertive Training," Journal of Abnormal Psychology, 76 (1970), 295-303.
- W. Mischel, Introduction to Personality (New York: Holt, Rinehart and Winston, Inc., 1971)
- M. Nakanishi, "Personality and Consumer Behavior: Extensions," Proceedings, Association for Consumer Research (1972), 61-65.
- M. Pfaff and S. Blivice, "Socioeconomic Correlates of Consumer Satisfaction and Activism," paper presented at the Research Symposium on Consumer Satisfaction/Dissatisfaction and Complaining Behavior, Indiana University, Bloomington, April 21-22, 1977.
- I. Stevenson and J. Wolpe, "Recovery from Sexual Deviations Through Overcoming Non-sexual Neurotic Responses," American Journal of Psychiatry, 116 (1960), 737-742.
- M. Wall, L.E. Dickey, and W. Talarzyk, "Predicting and Profiling Consumer Satisfaction and Propensity to Complain," paper presented at the Research Symposium on Consumer Satisfaction/Dissatisfaction and Complaining Behavior, Indiana University, Bloomington, April 21-22, 1977.
- R.H. Warland, R.O. Hermann, and J. Willits, "Dissatisfied Consumers: Who Gets Upset and Who Takes Action," Journal of Consumer Affairs, 9 (1975), 148-163.
- B. Weinman, P. Gelbart, M. Wallace, and M. Post, "Inducing Assertive Behavior in Chronic Schizophrenics: A Comparison of Socioenvironmental Desensitization, and Relaxation Therapies," Journal of Consulting Clinical Psychology, 39:2 (1972), 246-252.
- J. Wolpe, The Practice of Behavior Therapy (New York: Pergamon Press, 1969)
- J. Zaichowsky and J. Liefeld, "Profiling Personality Traits of Complaint Letter Writers," paper presented at the Research Symposium on Consumer Satisfaction/Dissatisfaction and Complaining Behavior, Indiana University, Bloomington, April 21-22, 1977.
- G. Zaltman, R.K. Srivastava, and R. Deshpande, "Perceptions of Unfair Marketing Practices: Consumerism Implications," in K. Hunt, ed., Advances in Consumer Research, 6 (1977), 247-252.
- W.G. Zikmund and S.J. Miller, "Consumerism Movements: An Examination of Attitudes and Behavior," Proceedings, Southern Marketing Association (1974), 5-10.



CORRELATIONS BETWEEN THREE  
INDICATORS OF BREADTH AND VARIATION IN  
SOURCES OF STIMULATION

Elizabeth C. Hirschman, New York University  
Melanie Wallendorf (student), University of Pittsburgh

ABSTRACT

The term 'stimulus variation' refers to variety in the types and sources of mental stimulation a person receives. This paper explores the relationships among three indicators of stimulus variation -- role accumulation, mass media utilization, and cosmopolitanism. Empirical evidence of correlation is provided and several conceptual issues are cited.

Introduction

The term 'stimulus variation' refers to variety in the types and sources of mental stimulation a person receives. Stimulus variation has been investigated in a non-systematic way in both psychological and sociological contexts, where it has been alternatively termed activity seeking (Fowler, 1976), novelty-seeking (Finger and Mook, 1971), information seeking (Thorelli, 1975), sensation-seeking (Zuckerman, Kolin, Pierce and Zoub, 1964), exploratory drive (Nissen, 1951), spontaneous alternation (Hosada, 1964), optimal stimulation level (Hebb and Thompson, 1954), incongruity seeking (Hunt, 1963), innovation proneness (Rogers and Shoemaker, 1971), exploratory urge (Cattell, 1957), venturesomeness (Robertson, 1970), and role accumulation (Sieber, 1974; Wallendorf, 1976; Zaltman, Wallendorf and Hirschman, 1977; Hirschman and Wallendorf, 1977). Depending on whether a psychological or a sociological approach is taken, engaging in a lifestyle which involves a high level of stimulus variation can be viewed as arising from an internal need or drive, or as being a condition which is socially induced.

Whatever its cause for a particular individual, the concept of stimulus variation is an important one for the study of consumer behavior for several reasons. First, the concept of stimulus variation is important for theories concerning consumer information processing and information overload (e.g., Scammon, 1977; Bauer, 1960; Hunt, 1963; Jacoby, Speller and Kohn, 1974a; Jacoby, Speller and Kohn 1974b; Katona and Mueller, 1955; Osgood and Tannenbaum, 1955; Ray, 1973; Russo, 1974), which posit that individuals may differ in their ability to process and utilize diverse bits of information. This may result in a situation in which, due to the diversity of the bits of information and individuals' differing abilities to handle stimulus variation, one person's information 'overload' constitutes another individual's 'underload'.

The concept of stimulus variation is also important for understanding how individuals choose among various mass media sources of information and stimulation. This may affect the type of media selection strategy which will be most effective in reaching a given target audience. The strategy for reaching target audiences containing large numbers of consumers whose chosen sources reflect stimulus variation will differ from the strategy for reaching audiences characterized by low stimulus variation.

It also appears logical that an individual could obtain stimulus variation in several different ways. For

example, it is possible that individuals who are desirous of a high level of stimulus variation (for either internal need-based reasons or for external socially-induced reasons) would have a lifestyle characterized by several types of stimulus variation. That is, the person who belongs to a varied set of groups or organizations and participates in a varied set of social activities might also be one who would expose himself to a varied set of mass media programs and publications and to geographically-dispersed stimuli. In other words, a person may obtain stimulus variation simultaneously in several different ways (e.g., interpersonally, impersonally, and geographically).

Purpose

The purpose of this exploratory research is to broaden research efforts already underway on the sociological aspects of the stimulus variation phenomenon to two additional areas: mass media utilization and cosmopolitanism. The area previously under investigation was that of role accumulation. A sociological construct, role accumulation refers to the number of groups to which an individual belongs.

Previous work has developed five propositions concerning role accumulation and its relationship to consumer behavior (Wallendorf, 1976). This work states the proposition that role accumulators are more likely to know about innovations earlier than others, and to adopt innovations earlier than others. It also makes the proposition that for role accumulators the time from awareness to adoption of an innovation will be shorter than for other individuals. Finally, this work makes the proposition that role accumulators will initiate and will be sought out for interpersonal communications about innovations more often than others. These relationships are shown in Figure 1. These relationships are presented here to indicate how one form of stimulus variation, role accumulation, may affect some consumption-related activities. These relationships are not, however, tested in the research reported here.

Instead, the purpose of this research is to broaden the applicability of the role accumulation concept and connect it with the more general phenomenon of stimulus variation. Three propositions were formulated and are given below.

1. The greater the degree of role accumulation, the greater the number of different types of mass media vehicles an individual will utilize.
2. The greater the degree of role accumulation, the greater the degree of cosmopolitanism.
3. The greater the number of different types of mass media vehicles utilized, the greater the degree of cosmopolitanism.

The relationships proposed in these three propositions are shown in Figure 2. These three propositions are the hypotheses tested in this research.

FIGURE 1  
Proposed Relationships Between Role Accumulation  
and Diffusion of Innovation Concepts

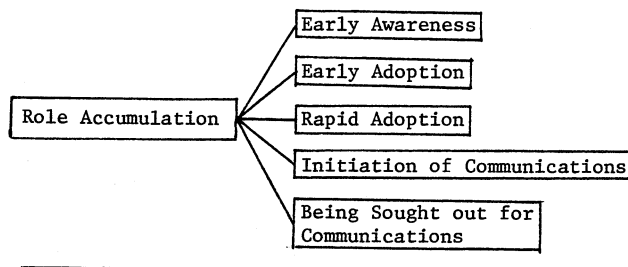
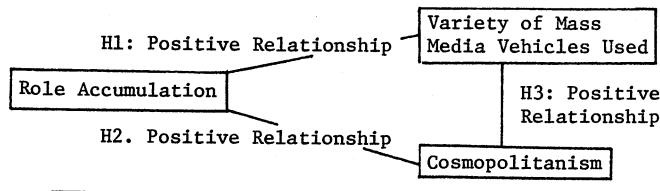


FIGURE 2  
Hypothesized Relationships Between Three  
Types of Stimulus Variation



#### Methodology

To test these three hypotheses, survey data from two research projects were utilized. The first survey consisted of 587 telephone interviews conducted with residents of the Birmingham, Alabama SMSA in August 1977. The second survey was composed of 504 telephone interviews conducted with residents of the Decatur, Illinois SMSA in December 1977. Both surveys were based on random-digit-dialed telephone interviews conducted by professional interviewers. Interviewing was conducted during the evening hours to assure an adequate representation of working persons. Sampling from each telephone exchange in the SMSA was weighted according to population, and the obtained sample was consistent with updated census data for that SMSA in all demographic variables studied (i.e., race, age, occupation, education, and income). During each interview, information was gathered on (1) the individual's group memberships and social activities, (2) mass media utilization, and (3) cosmopolitanism.

#### Group Memberships and Social Activities

Respondents were asked to name the groups to which they belonged in each of five categories: business, community, social, recreational and religious. They were also asked to give the extent to which they participated in several recreational and entertainment activities.

#### Mass Media Utilization

Respondents were asked to name the mass media vehicles they regularly used as sources of information. Due to questionnaire length constraints, these were limited to radio, newspaper and television vehicles in the Decatur survey, and newspaper and magazine vehicles in the Birmingham survey.

#### Cosmopolitanism

Cosmopolitanism refers to the extent to which an individual is involved in affairs, experiences, and relationships outside the geographically local social system. In the literature dealing with cosmopolitanism, it has been variously measured by finding out how often the individual travels to other cities, how much

exposure the person has to cosmopolite communication channels such as national organization newsletters, how many people the individual knows outside the local area, and how geographically dispersed these people are (Rogers and Shoemaker, 1971).

In this study, respondents were asked to what extent they went travelling or sightseeing. This question actually measures only one component of cosmopolitanism, that dealing with physical ties to other areas. Other components dealing with relational or organizational ties to other areas were not measured. For the purposes of this study, however, the question on travelling and sightseeing was deemed sufficient. This question was included in the larger set of questions concerning participation in various social activities.

#### Analysis

#### Operationalizing Role Accumulation

Two measures of role accumulation were utilized. Because a measure of role accumulation should represent the variety of group memberships and social activities engaged in by the individual, one way of operationalizing the concept is to sum the number of different types of groups and social activities in which the individual participates. Therefore one point was given for each type of organization to which the individual belonged. For example, if a respondent belonged to 2 business groups, 3 community groups, and 1 religious group, the group membership score would be 3 (one point for each type of group for which the individual reported membership). This measure is designed to indicate variety rather than merely sheer number of group memberships.

The second component of this role accumulation index was based on the number of different social activities in which the individual participated. A score of one was given for each of the seven social activities in which the individual participated very often, moderately, or not too often. (The sightseeing and travelling item was not included in this index). For example, if a respondent indicated that he or she participated in tennis very often, home entertaining not too often, and never in the other activities, the social activities component of the role accumulation index would be a score of two. In summary, the first measure of role accumulation was number of different types of groups to which the individual belonged plus the number of different social activities in which the individual participated. This index was created for each individual, and the results are given in Table 1 for both the Birmingham and Decatur data. As can be seen from these figures, the incidence of role accumulation using this index generally follows a normal distribution, but is slightly skewed to the right. The data in Table 1 further indicate that the distribution of this index of role accumulation is generally consistent between Decatur and Birmingham.

There are some problems inherent in this summed operationalization of role accumulation, because redundancy may be present among some of the variables. For example, by summing together an individual's recreational group memberships and his or her participation in tennis and golf there may be some double-counting. Secondly, it is possible that involvement in certain types of clubs or activities may be negatively associated with participation in others. For example, people who are avid concert-goers may be averse to attending spectator sporting events. By summing together all activities and group memberships the occurrence of such negative correlations may be masked.

To overcome these two potential shortcomings of the summed measure of role accumulation, a second operationalization was also used. This consisted of factor

analyzing the group memberships and scores on the seven social activities. Such a procedure should serve to reduce any redundancy and extract distinct, orthogonal factors underlying the data. To accomplish this a Principal Factors Analysis combined with a Varimax rotational procedure was employed on the group membership and social activity variables in both cities. The results are given in Table 2A. In both cities three factors were obtained which had eigenvalues substantially greater than one. A total of 44.3% of variance was explained by the three factors in Decatur, and 43.4% total variance was explained in Birmingham.

Two of the factors obtained in each city appeared to be fairly consistent. Factor 1 in Decatur and Factor 3 in Birmingham both loaded very highly on recreational group memberships, golfing, tennis playing, movie attendance and attendance of spectator sporting events. Factor 2 in Decatur and Factor 2 in Birmingham also appeared to be analogous. These loaded highly on membership in business, social, community, recreational and religious group memberships in both cities. The remaining factors in each city -- Factor 3 in Decatur and Factor 1 in Birmingham -- possess some similarities, but do not appear to be as analogous as the other two sets of factors. For example, both factors are consistent in their high loadings for concert attendance and home entertaining, yet are inconsistently loaded on spectator sports attendance, movie attendance, bridge club participation and community group membership.

The scores an individual had on each factor were added together to create a composite principal factor index of role accumulation. The distributions obtained for this principal factor index of role accumulation are given in Table 2B.

TABLE 2A  
Role Accumulation Factor Indexes

Factor Analysis: Decatur

	<u>Factor 1</u>	<u>Factor 2</u>	<u>Factor 3</u>
Business Organizations	0.09310	0.47170*	-0.04057
Community Organizations	-0.04162	0.51439*	0.37447*
Social Organizations	0.02229	0.68417*	0.19075
Religious Organizations	-0.00650	0.25333*	0.21276
Recreational Organizations	0.45610*	0.37457*	-0.02446
Camping	0.21881	0.02476	-0.20336
Golfing	0.38841*	0.07292	0.35406*
Tennis	0.62633*	0.00611	0.19186
Home Entertainment	0.18504	0.07751	0.38849*
Bridge	0.08075	0.06339	0.38536*
Concert	0.16369	0.16788	0.47718*
Movies	0.54186*	-0.07831	0.47718*
Spectator Sports Events	0.40977*	0.06257	0.17644

Factor Analysis: Birmingham

	<u>Factor 1</u>	<u>Factor 2</u>	<u>Factor 3</u>
Business Organizations	0.09138	0.21434	0.07993
Community Organizations	0.09701	0.66760*	0.14929
Social Organizations	0.06009	0.52985*	0.03399
Religious Organizations	0.04652	0.49949*	-0.05924
Recreational Organizations	0.01851	0.42055*	0.41088*
Golfing	0.12439	0.05485	0.41413*
Tennis	0.26796*	-0.00447	0.56414*
Home Entertaining	0.49637*	0.21759	0.14656
Bridge	0.07668	0.04181	0.02838
Concert	0.56730*	0.22502	0.08951
Movies	0.63024*	-0.12542	0.25253*
Spectator Sports Events	0.45521*	0.06552	0.31152*

\* Indicates factor loading higher than 0.25

TABLE 1  
Role Accumulation:  
Summed Operationalization Distributions

<u>DECATUR</u>			<u>BIRMINGHAM</u>		
<u>SCORE</u>	<u>N</u>	<u>(%)</u>	<u>SCORE</u>	<u>N</u>	<u>(%)</u>
0.	11	2.2	0.	16	3.7
1.	33	6.5	1.	32	5.5
2.	43	8.5	2.	35	6.0
3.	74	14.7	3.	49	8.3
4.	78	15.5	4.	68	11.6
5.	77	15.3	5.	83	14.1
6.	60	11.9	6.	101	17.2
7.	50	9.9	7.	82	14.0
8.	36	7.1	8.	48	8.2
9.	26	5.2	9.	36	6.1
10.	12	2.4	10.	24	4.1
11.	2	0.4	11.	9	1.6
12.	2	0.4	12.	4	0.7
TOTAL	502	100.0	TOTAL	587	100.0

Mean = 4.83                      Mean = 5.43  
Standard Deviation = 2.45      Standard Deviation = 2.6

TABLE 2B  
Role Accumulation:  
Factor Index Operationalization Distributions

<u>DECATUR</u>			<u>BIRMINGHAM</u>		
(scores collapsed into integer categories)			(scores collapsed into integer categories)		
<u>SCORE</u>	<u>N</u>	<u>(%)</u>	<u>SCORE</u>	<u>N</u>	<u>(%)</u>
-3	2	0.4	-3	28	4.8
-2	96	18.8	-2	101	17.9
-1	164	32.4	-1	124	21.1
0	84	16.7	0	77	13.4
1	66	13.0	1	62	10.6
2	43	8.6	2	35	5.9
3	22	4.5	3	18	3.1
4	14	2.6		587	100.0
5	5	1.0			
6	3	0.6			
7	3	0.6			
8	2	0.4			
9	1	0.2			
	502	100.0			

Mean = .063

Mean = -.007

Standard Deviation = 2.12

Standard deviation = 1.75

### Operationalizing Media Utilization

In each media category the objective was to obtain a measure of the breadth of an individual's media exposure rather than the depth of exposure. Thus, the number of different types of vehicles a person was exposed to was more relevant than the sheer amount of media vehicle exposure. Therefore vehicles in each medium (i.e., radio, television, magazine, and newspaper) were subdivided into several categories which appeared substantively different. An index of mass media utilization breadth was then constructed by summing the number of different categories a person was exposed to within each medium.

Categories were arbitrarily defined for each medium. Newspapers were grouped into four categories: (1) local newspapers, (2) national newspapers, (3) financial newspapers (e.g., Wall Street Journal), and (4) special interest newspapers (e.g., Village Voice). Magazines were divided into five groups based on information content. The categories were: (1) news magazines (e.g., Time), (2) fashion magazines (e.g., Vogue), (3) home magazines (e.g., Ladies Home Journal), (4) men's magazines (e.g., Playboy), and (5) special interest magazines (e.g., Photography, Yachting, and Ebony).

Radio stations were separated into three groups according to programming format: (1) rock/soul, (2) easy listening, and (3) special interest (e.g., classical music, all news, religious). Television programs were initially divided into fifteen categories which included situation comedies, daytime soap operas, night-time soap operas, drama, serials, sporting events, police shows, talk shows, movies, local news programs, national news programs, variety shows, miscellaneous PBS programs, religious shows, and specials. However, realizing that this was a rather large set of categories, a reduced set of television program types was also developed. This reduced set of television programs was composed of six categories: (1) local news programs, (2) national news programs, (3) specials, (4) movies, (5) talk shows, and (6) PBS programming. These were selected because they were judged to be more "stimulation rich" than other types of television programming, and therefore to perhaps better represent the concept of stimulus variation.

### Mass Media Utilization Index

Decatur. In Decatur respondent utilization of newspapers, radio stations and television programming was measured. To combine these various media vehicles into an index of media utilization which would reflect the concept of stimulus variation, a summing procedure was employed. That is, the number of different categories of each medium (i.e., radio, newspapers, and television) in which the individual utilized at least one vehicle were summed to create an index for that individual.

To further refine the index a secondary analysis was made of "special interest" categories for newspaper readership and radio station listenership for each individual. Where the specialized vehicles in each medium appeared to represent no overlap, (e.g., a classical music station and an all news station), the individual's index was increased by the number of different special interest categories represented.

Finally, a second media utilization index was created for each individual which included the radio station and newspaper categories used in the first index, but the reduced set of "stimulation rich" television program types.

Birmingham. In Birmingham respondent utilization of the two print media, newspapers and magazines, was measured.

Similar to the approach used in Decatur, a summed index was created for each individual by adding together the number of different types of magazines and newspapers utilized. Also an analysis was made of the specific vehicles contained in the "special interest" categories for each person and, where the vehicles were not redundant, the respondent's index was incremented.

The distributions for the media utilization indices created in Decatur and Birmingham are shown in Table 3. In summary, three media utilization indices were created: (1) the use of print media in the Birmingham sample; (2) the use of the broadcast media and newspapers in the Decatur sample with a large set of categories for television programming; and (3) the use of the broadcast media and newspapers in the Decatur sample but with a reduced set of categories for television programming.

### Operationalizing Cosmopolitanism

Cosmopolitanism was operationalized in this research as being the frequency with which the individual was exposed to sources of external stimulation through sightseeing and travelling. The distribution of responses to this measure is given in Table 4 for both Decatur and Birmingham.

### Correlational Analysis to Test Hypothesis One: Role Accumulation and Media Utilization

To measure the degree of correspondence between role accumulation and media utilization, correlational analyses were performed using the two operationalizations of role accumulation, (i.e., the summed index and the principal factors index).

The results for the summed index of role accumulation are given in Table 5. As can be seen, the correlations between the role accumulation summed index and media utilization were consistently positive and statistically significant at the  $p = .003$  level or beyond. However, there were some variations in the degree of correlation. The weakest relationship was found to exist between the "full set" media utilization measure (i.e., that containing all television program types) and role accumulation in the Decatur data. An increased level of correlation was obtained when television program types included in the media utilization measure were "reduced" to those judged to be "information rich" in terms of stimulus variation. By using this reduced set, the correlation increased from .124 to .134 and the significance level from .003 to .001.

An even higher level of correlation and statistical significance was found in Birmingham when the role accumulation summed index was correlated with a media utilization index created from the two print media. The  $R$  in this instance was .389 which was significant at  $p < .001$ .

This finding appears to indicate that role accumulation, at least as measured by summing across group memberships and activities, is more highly related to stimulus variation for print media, alone, than for broadcast and print media combined. One explanation for this is that perhaps a greater inherent variety of stimulation is found in magazines and newspapers than in television or radio programming. The validity of this conjecture is inadequately addressed in this research, but, it would seem to be an area worthy of further inquiry.

Table 5 also shows the results obtained from correlating the principal factors role accumulation index with media utilization. In Decatur, where the two media utilization indexes were composed of newspaper, radio and television usage, correlations were positive but lower than obtained in Birmingham. The lowest correlation,  $R = .083$  was again found between the full measure of media

utilization (which included all types of television programs) and the role accumulation principal factors index. The correlation was increased to .113 when the role accumulation principal factor index was correlated with the reduced set measure of media utilization (which included only six types of television programming).

As was found with the summed measure of role accumulation, the role accumulation principal factor index appears to be more highly associated with stimulus variation in print media, alone, than with broadcast and print media combined. Whether this is due to an inherently greater availability of variety in the stimulation obtainable from print media or to some other, unmeasured variable remains unknown.

Correlational Analysis to Test Hypothesis Two: Role Accumulation and Cosmopolitanism

The results obtained from correlating role accumulation with cosmopolitanism are given in Table 5 for both the summed index of role accumulation and the principal factor index of role accumulation. As is shown in Table 5, a high degree of correlation exists between the summed index of role accumulation and cosmopolitanism. In Decatur that R value was .357, while in Birmingham an R value of .522 was found. Both of these correlations were significant at beyond the .001 level.

Similar results were obtained when the principal factor index representing role accumulation was correlated with cosmopolitanism. In Decatur the correlation was .309 and was significant at the .001 level. The Birmingham data indicated an even higher pattern of correlation between the role accumulation principal factor index and cosmopolitanism. The R value was .465 and the significance level .0001.

Correlational Analysis to Test Hypothesis Three: Media Utilization and Cosmopolitanism

To test the third hypothesis that persons using a wider variety of mass media vehicles would exhibit a greater degree of cosmopolitanism, the correlation between these two variables was computed. As is shown in Table 5 in both Decatur and Birmingham the correlations were found to be positive and statistically significant at the  $p = .001$  level and beyond.

It was found that the breadth of the utilization of print media alone (Birmingham) had a higher correlation with cosmopolitanism than did utilization of print and broadcast media combined (Decatur). This finding was reminiscent of the results obtained when media utilization was correlated with role accumulation. The reasoning advanced to explain that finding may be equally applicable here. That is, the stimulus variation provided by television and radio programming may not be as great as that provided by print media. It is also useful to note that when the "full set" of television programs included in the Decatur media utilization index was "reduced" to those shows judged to be richer in stimulation, a large improvement in correlation, rivaling that in Birmingham, was found.

TABLE 3  
Media Utilization Index Distributions

DECATUR (Full Set)			DECATUR (Reduced Set)		
SCORE	N	(%)	SCORE	N	(%)
0.	3	0.6	0.	13	2.2
1.	21	4.2	1.	50	11.5
2.	62	12.3	2.	149	29.6
3.	100	19.8	3.	148	29.4
4.	134	26.6	4.	92	18.3
5.	99	19.6	5.	34	6.7
6.	58	11.5	6.	7	1.4
7.	23	4.6	7.	5	1.0
8.	1	0.2		504	100.0
9.	2	0.4			
10.	1	0.2			
	504	100.0			

Mean = 4.01, S.D. = 1.56      Mean = 2.81, S.D. = 1.89

BIRMINGHAM

SCORE	N	(%)
0.	129	22.0
1.	177	30.2
2.	103	17.5
3.	80	13.6
4.	92	18.3
5.	20	3.4
6.	10	1.7
7.	1	0.2
	587	100.0

Mean = 1.80, S.D. = 1.54

TABLE 4  
Cosmopolitanism Index Distributions

	N	(%)		N	(%)
Very often	116	23.0	Very often	109	18.6
Moderately	185	36.7	Moderately	180	30.7
Not too often	148	29.4	Not too often	185	31.6
Never	55	10.9	Never	112	19.1
	504	100.0		586	100.0

Mean = 2.72, S.D. = .94      Mean = 2.48, S.D. = 1.00

TABLE 5  
Correlational Analyses

(Role accumulation Summed Index  
Correlated With Media Utilization)

DECATUR (F)*	DECATUR (R)*	BIRMINGHAM
R = .124	R = .134	R = .389
P < .003	P < .001	P < .001

(Role Accumulation Principal Factor Index  
Correlated With Media Utilization)

DECATUR (F)	DECATUR (R)	BIRMINGHAM
R = .083	R = .113	R = .357
P < .03	P < .005	P < .001

TABLE 5  
(continued)

(Role Accumulation Summed Index  
Correlated With Cosmopolitanism)

DECATUR	BIRMINGHAM
R = .309	R = .465
P < .001	P < .0001

(Media Utilization Correlated With Cosmopolitanism)

DECATUR (F)	DECATUR (R)	BIRMINGHAM
R = .153	R = .265	R = .278
P < .001	P < .001	P < .001

\* F = Full set, R = Reduced set

#### Summary and Implications

The consistently positive and statistically significant correlations among role accumulation, variety in media utilization, and cosmopolitanism found in both cities provide support that these three types of stimulus variation are related. The finding that stimulus variation may be generalized across several areas of an individual's behavior and particularly those relating to information-seeking activities, could have implications for the analysis and understanding of many consumer behavior phenomena.

To discuss these implications, focus will first be placed upon the individual acting within a social system and than upon social systems, generally. First, if stimulus variation activities are normally distributed within a population, and the results reported here appear to support this contention, then it follows that the majority of people will be characterized by a moderate amount of stimulus variation, while others will exhibit higher or lower amounts.

It is the individuals positioned at either end of the continuum who are perhaps most important to consumer behavior research and about whom several conjectures can be made. First, those persons exhibiting a below average amount of stimulus variation may have a reduced ability (psychologically, socially, and/or physiologically induced) to obtain and process information. This below average amount of information seeking and processing may reduce their ability to adequately evaluate relevant alternatives in many consumption activities and generally may impair their functioning in a complex, consumption-oriented society. One would expect people with low amounts of stimulus variation to be largely habitual in their purchasing behavior and possibly unable or unwilling to alter this behavior, even in the face of changing circumstances.

At the other end of the spectrum may be found those persons who experience a high amount of stimulus variation. From a broad-based consumer behavior perspective these individuals would be the consummate consumers -- continuously being exposed to information about new products, new knowledge and new experiences. They are of particular interest because they may be the people most likely to be reached by any mass media information flow, whether it be commercially sponsored, public interest or otherwise.

Their value to an increased understanding of consumer behavior, however, arises not only from their heightened rate of product consumption and information acquisition activities, but from their above average rate of information distribution, as well. Individuals who have a high amount of stimulus variation may play a central role in the diffusion process. Due to their diverse and extensive contact with new information, such people may be the first to learn about new products, and among the earliest to adopt them. Additionally, because of their widespread network of social activities and group memberships, they could serve as disseminators of information and ideas about new products. These implications can be illustrated by mentally connecting Figures 1 and 2.

Shifting focus to the influence which stimulus variation may have on social consumption patterns generally, one obvious issue is the longitudinal distribution of stimulus variation within the population. Sociologist Peter Blau (1975) has hypothesized that a process of increasing multiform heterogeneity is endemic to most Western societies, with the United States serving as the archetype. Blau's thesis asserts that societies are not only becoming more complex, but that people's networks of social contracts are becoming wider and more fractionated. That is, people are becoming exposed to an increasing variety of stimuli from diverse sources -- the media, friends, business associates, educational institutions and special interest groups.

From a sociological point of view one of the key concerns is the effects these increasing levels of multiform heterogeneity may have on society. The most constructive scenario, perhaps, would be the occurrence of a gradual and nondisruptive movement by the entire population toward adaptation to greater levels of social complexity. That is, the distribution of stimulus variation within the population would maintain its normal shape but be shifted steadily rightward.

Another, more disconcerting possibility is that instead of gradually adapting themselves to higher levels of stimulation, a large portion of the population would begin to increasingly reject change and move instead toward desiring lower levels of stimulus variation. If this situation holds, and some recent research supporting an apparent resurgence in ethnicity suggests that it may (Glaser and Moynihan, 1975), then society may become bipolar with respect to stimulus variation. The distribution mean would remain constant, but one portion of the population would withdraw from stimulus variation, while another portion moved forward to embrace it, giving the distribution a bimodal character.

If the first of these scenarios is the more valid one, what alterations can be expected in consumption behavior? First, a primary result may be an increase in the rapidity with which new products (broadly defined) diffuse. A higher level of mass media monitoring and a broadened pattern of interpersonal contacts should prove highly conducive to new product adoption and diffusion. Closely related to this, however, is the notion that there would be a correspondingly shorter life cycle for a given product, as it would be more readily discarded in favor of something novel.

Conversely, if the second scenario of a bipolar society should become a reality, its implications for consumer behavior are also significant. First, marked differences between large population segments in their willingness and ability to process information may effectively create a two-tier consumption system. Within one tier high-stimulus-variation individuals may manifest the type of behavior described earlier. That is, new products would be learned about, adopted and discarded rapidly. Within the second tier, the diffusion

of new products would be retarded as people maintained habitual consumption patterns and were unwilling or unable to seek out new information or to adopt novel consumption alternatives.

The implications of these two possible societal scenarios are large not only for consumer behaviorists but for political scientists and sociologists as well. Although longitudinal analyses are required for studying such a phenomenon, some clues as to the relative validity of each scenario may be gained by studying societies in different stages of development and looking for shifts toward bipolarization or increased over-all levels of stimulus variation.

The research reported here is unable to address these larger issues because of the similarities in size and social composition of the two social systems studied. However, the genuine value such analyses would have should serve as adequate impetus to researchers for extending investigation beyond the preliminary boundaries established here.

#### References

- Raymond A. Bauer, "Consumer Behavior as Risk Taking, in R.S. Hancock (Ed.), Dynamic Marketing for a Changing World, Proceedings of the 43rd National Conference of the American Marketing Association (1960), 389-440.
- Peter M. Blau, "Presidential Address: Parameters of Social Structure," American Sociological Review, Vol. 39 (October 1975), 615-635.
- R.B. Cattell, Personality and Motivation Structure and Measurement (New York: Harcourt, Brace & World, 1975).
- Frank W. Finger and Douglas G. Mook, "Basic Drives," Annual Review of Psychology, Vol. 22 (1971), 8-9.
- Harry Fowler, "Satiation and Curiosity: Constructs for a Drive and Incentive-Motivational Theory of Exploration," in K.W. Spence and J.T. Spence (Eds.), Psychology of Learning and Motivation, Vol 1 (New York: Academic Press, 1967), 157-277.
- Nathan Glaser and Daniel P. Moynihan (Eds.), Ethnicity: Theory and Experience (Cambridge, MA: Harvard University Press, 1975).
- D.O. Hebb and W.R. Thompson, "The Social Significance of Animal Studies," in G. Lindsey (Ed.), Handbook of Social Psychology (Reading, MA: Addison-Wesley, 1954), 551-2.
- Elizabeth C. Hirschman and Melanie Wallendorf, "Role Accumulation and Early Adoption," Working Paper #260, Graduate School of Business, University of Pittsburgh, Pittsburgh, PA (1977).
- K. Hosada, "Curiosity and Spontaneous Alternation," Japanese Journal of Psychology, Vol. 35, 3 (1964), 126-32.
- J.M.V. Hunt, "Motivation Inherent in Information Processing and Action," in O.J. Harvey (Ed.), Motivation and Social Interaction - Cognitive Determinants. (New York: Ronald Press, 1963), 35-94.
- Jacob Jacoby, Donald E. Speller and Carol A. Kohn, "Brand Choice Behavior as a Function of Information Load," Journal of Marketing Research, Vol II (1974), 63-9.
- Jacob Jacoby, Donald E. Speller and Carol A. Kohn, "Brand Choice Behavior as a Function of Information Load: Replication and Extension," The Journal of Consumer Research, Vol.1 (1974), 33-42.
- G. Katona and E. Mueller, "A Study of Purchase Decision," in L.H. Clark (Ed.), Consumer Behavior: The Dynamics of Consumer Reaction (New York: New York University Press, 1955), 30-87.
- Henry W. Nissen, "Phylogenetic Comparison," in S.S. Stevens (Ed.), Experimental Psychology (New York: John Wiley & Sons, 1951).
- C.E. Osgood and P.H. Tannenbaum, "The Principle of Congruity in the Prediction of Attitude Change," Psychological Review, Vol. 62 (1955), 42-45.
- M. Ray, "Psychological Theories and Interpretation of Learning," in S. Ward and T.S. Robertson (Ed.), Consumer Behavior: Theoretical Sources, (Englewood Cliffs, NJ: Prentice-Hall, 1973), 45-117.
- T.S. Robertson, Consumer Behavior, (Glenview, IL: Scott Foreman Co., 1970).
- Everett M. Rogers and E. Floyd Shoemaker, Communication of Innovations, (New York: The Free Press, 1971).
- J.E. Russo, "More Information is Better: A Reevaluation of Jacoby, Speller and Kohn," The Journal of Consumer Research, Vol. 1 (1974), 68-72.
- Debra L. Scammon, "Information Load and Consumers," Journal of Consumer Research (December, 1977), 148-155.
- Sam Sieber, "Toward a Theory of Role Accumulation," American Sociological Review. Vol. 39 (August, 1974), 567-578.
- Hans B. Thorelli, Helmut Becker and Jack Engledow, The Information Seekers: An International Study of Consumer Information and Advertising Image (Cambridge, MA: Ballinger, 1975).
- Melanie Wallendorf, "Exploring the Impact of Role Accumulation on the Diffusion of Innovations," Working Paper, Graduate School of Business, University of Pittsburgh, Pittsburgh, PA (1977).
- Gerald Zaltman, Melanie Wallendorf and Elizabeth C. Hirschman, "Role Accumulators Versus Innovators, Opinion Leaders and Gatekeepers as Important Groups in Consumer Research," Working Paper, Graduate School of Business, University of Pittsburgh, Pittsburgh, PA (1977).
- M. Zuckerman, E.A. Kolin, L. Price and I. Zoub, "Development of a Sensation-Seeking Scale," Journal of Consulting Psychology, Vol. 38 (1964).

## DOGMATISM AND INNOVATION: A SITUATIONAL PERSPECTIVE

Kenneth A. Coney, Arizona State University  
Robert R. Harmon, Portland State University

### Abstract

The influence of the gift giving situation on the well established dogmatism-innovation relationship is demonstrated. High dogmatics in a gift giving situation were found to be more innovative than low dogmatics in the same situation. More importantly, the direction of the dogmatism-innovation relationship under the gift giving situation was reversed.

### Introduction

Generally, researchers in consumer behavior have found a disappointing lack of predictive power on the part of personality variables relative to purchase and choice phenomena. As Kassarian (1971) points out, however, it is almost too much to expect specific variables alone to explain a great deal of the variance in complex behavioral phenomena. To some extent, the development of psychographics has been a response to this need for a combination of variables to better understand and predict behavior (Belk 1975a). Recently, however, Belk (1975b) has provided a potentially more powerful explanation and conceptualization of why there is such a lack of predictive relationships between personality variables (or any other potentially predictive variables) and behavior have been found. Although Belk's view is certainly not new (Ward and Roberson 1973, Lavidge 1966) it is more productive than earlier discussions as Belk provides a taxonomy of situations for research consideration.

Specifically, Belk (1975b) defines five groups of situational characteristics: 1) Physical Surroundings, 2) Social Surroundings, 3) Temporal Perspective, 4) Task Definition, and 5) Antecedent States. Any of these situations in interaction with a specific object and person can influence subsequent behavior. Although there is still a question as to the best way to define a situation, particularly in terms of degree of generality (Barker and Wicker 1975, Belk 1975b), Belk's taxonomy seems reasonable to use as a starting point to determine the predictive effect of personality and situation on various behavioral phenomena.

Of particular interest to this paper is Belk's (1975b) task definition characteristic of a situation. He defines these features to be:

... an intent or requirement to select, shop for, or obtain information about a general or specific purchase. In addition, task may reflect different buyer and user roles anticipated by the individual. For instance, a person shopping for a small appliance as a wedding gift for a friend is in a different situation than he would be in shopping for a small appliance for personal use (p. 159).

The task definition of gift giving as described above would seem to be an important aspect of consumer behavior. Belk (1976) states that the gift giver must make inferences about the recipient's tastes, needs, desires, possible reactions and the gift-recipient relationship in order to complete the gift selection process (p. 155). Furthermore, the task requirement of the gift giver situation could conceivably have a much different impact on purchase behavior than the task require-

ment of a purchase for personal consumption situation. Indeed, if the changing of the task requirement from that of purchase for personal use to that of purchase for the purpose of giving a gift to another, resulted in a demonstrable and systematic effect on behavior, then this would lend support to Belk's inclusion of task definition in his situational taxonomy.

### Situation, Personality, and Behavior

As innovative product choice is a real and important behavioral phenomenon to marketers, it seems important to determine if such behavior may be influenced in part by certain situations. Since the dogmatism-innovation relationship has already been established in the purchase for personal consumption situation it appears reasonable to extend the analysis of the relationship into another aspect of the task situation, in this case gift giving. Gift giving was chosen as a potentially important task situation as it represents a large proportion of purchase in many product classes (Belk 1976, Caron and Ward 1975). Also, it would seem that gift purchasing might have a very real impact on innovative products, where consumers may frequently want to be seen as giving something "new" and "different."

The relationship between dogmatism and innovation has been demonstrated by Jacoby (1971) and replicated by Coney (1972). Jacoby found that dogmatism and innovation were inversely related when individuals intended to buy for themselves ( $r = .36, p < .01$ ). Unfortunately, Jacoby's methodology did not account for the task situation variables and actually succeeded in lumping them together. He instructed his subjects: "If you do not use a set of products, buy for someone else." Therefore, Jacoby's study was compounding the effects of the situational influence of buying for oneself vs. buying for someone else. Not controlling for this dichotomous variable assumes that the situation does not make any difference in the dogmatism-innovation relationship.

Coney, using male subjects instead of female subjects found an even stronger inverse relationship between dogmatism and innovation ( $r = -.474, p < .001$ ). Coney's results may have been stronger because he did not confuse the dichotomous task requirement situation. Subjects were instructed to purchase for themselves. Therefore, his study only dealt with one half of the task situation dichotomy of interest to this present study. It is felt that the inclusion of the "buy for another" gift giving condition to the "buy for self" situational condition that was implicit in the Coney study and apparently dominant in the Jacoby study, would provide a good test of the predictive power of situation, object and person on behavior.

The intent of this study is to test the general null hypothesis that varying the task requirement of a purchase situation will have no impact on the dogmatism-innovation relationship. Rejection of this hypothesis should indicate that the specific situation has a direct effect on the strength and/or direction of the well-established dogmatism-innovation relationship. Such a finding would have direct implications for the inclusion of the situation in research on consumer purchasing pro-



cesses. Again, the thrust of this research effort is to emphasize the potential effect of specific situational variables on general relationships between marketing related phenomena.

### The Methodology

The same basic methodology used by Jacoby (1971) and Coney (1972) was followed. However, unlike the preceding studies, a dichotomous situational variable was added. The conditions are defined as the task requirement to (1) buy a product for one's personal use, and (2) buy a product as a gift for a friend. The friend was defined for subjects as someone close enough to the subject that they would realistically consider buying them a gift in the product categories shown. The same method of defining and determining innovative products was used in this study as in the previous two in order to replicate as closely as possible the basic methodology used to establish the inverse dogmatism-innovation relationship. The subjects were 110 unmarried male volunteers ranging in age from 19 to 28 years of age. All were business administration majors at Arizona State University.

The twelve product categories used in the study are listed in Table 1. These twelve were chosen from a larger list by a pretest group of 40 subjects similar to the general subject group. The products were determined prior to data collection to be ones that were familiar to and frequently purchased by the subject group. Every effort was made to provide subjects with product category choices that they currently purchase or would realistically be purchasing in the foreseeable future for themselves or give as a gift to a friend.

TABLE 1  
PRODUCT INNOVATION PRETEST

Product	Number Choosing Correct Innovation <sup>a</sup>	Chi-Square <sup>b</sup>
Watches	36	67.6
Belts	28	32.4
Cameras	26	25.6
Tennis Rackets	38	78.4
Wallets	25	22.5
Clock Radios	32	48.5
Beer	37	72.9
Briefcases	25	22.5
Coffee Makers	39	84.1
Jeans	36	67.6
Calculators	23	16.9
Casual Shoes	38	78.4

<sup>a</sup>N = 40.

<sup>b</sup>Significant at  $p < .001$ .

As subjects would be asked to choose one brand from a group of four brands in a product category, it was necessary to insure that the one choice defined as innovative (functional, fashion, or new brand) by the researchers was in fact seen as being innovative by subjects. The pretest group was shown the four brands in each category and asked to pick the one that they thought was an innovation. One could expect a chance innovation response rate of 10 when using a sample size of 40. As can be seen from Table 1, all product choices designated to be innovative were in fact considered to be so by the subjects beyond the .001 level of significance.

Subjects were administered the Rokeach Dogmatism Scale and divided into high and low dogmatic groups (see Table 2). Subjects in the experimental conditions were instructed to choose a product from each of the twelve product categories that they would be most likely to purchase (1) for themselves, or (2) as a gift for a male friend. The four pictures in each category were randomly mounted and could be viewed at the same time. All pictures were in color and matched on factors such as branded-unbranded, size and price.

The dependent measure used was purchase intention. While it would have been preferable to have an actual measure of behavior, given the limitations of the experiment, buying intentions seems to be an acceptable proxy. Fishbein (1975), for example, argues that a good predictor of a person's behavior is his/her intentions to perform the behavior irrespective of the nature of the behavioral criterion.

### Results

Subjects in each situational category were partitioned into two groups, based on their median dogmatism score. The 2 X 2 factorial design with mean innovation scores is shown in Figure 1 (a) and (b).

FIGURE 1

#### MEAN INNOVATION SCORES

(a)

Situation	Personality	
	High Dogmatic	Low Dogmatic
Buy for Self	3.20	4.12
Gift for Another	4.50	3.40

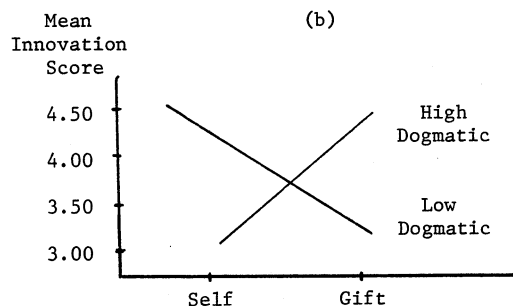


Table 2 shows the mean dogmatism score for high and low dogmatics for each situation. The expected inverse relationship was again found in the Buy for Self situation, however, a direct positive relationship was found in the Gift situation.

TABLE 2

MEAN DOGMATISM AND INNOVATION SCORE DIFFERENCES

Buy for Self Condition		
	Low Dogmatics (n = 25)	High Dogmatics (n = 25)
Dogmatism	$\bar{X} = 116.24$ $\sigma = 11.50$	$\bar{X} = 150.76$ $\sigma = 13.47$
	9.75 <sup>a</sup> p < .001	
Innovation	$\bar{X} = 4.12$ $\sigma = 1.42$	$\bar{X} = 3.20$ $\sigma = 1.45$
	2.57 <sup>a</sup> p < .01	
Simple r = -.41154, p < .001		
Gift Buying Condition		
	Low Dogmatics (n = 30)	High Dogmatics (n = 30)
Dogmatism	$\bar{X} = 112.77$ $\sigma = 14.85$	$\bar{X} = 156.77$ $\sigma = 14.35$
	11.67 <sup>a</sup> p < .001	
Innovation	$\bar{X} = 3.40$ $\sigma = 1.23$	$\bar{X} = 4.50$ $\sigma = 1.82$
	2.74 <sup>a</sup> p < .01	
Simple r = .10704, N.S.		

<sup>a</sup> One-tailed t-tests.

Table 3 shows the results of the ANOVA. Main effects for situation and dogmatism are not significant. However, the 2-way interaction is significant at p < .001.

These results are similar to earlier situational findings that demonstrated the dominance of interactions over main effects in contributions to variance (Belk 1974). In this instance, the significant interaction term suggests that innovative buying intentions are related to dogmatism and to the specific situational condition. The non-significant main effects can be interpreted as dogmatism and situation, when considered alone, cannot lead us to any general rules as to which level of either has the greatest impact on innovation.

In the buying for self situation the relationship between dogmatism and innovation explains 17.0% of the variance accounted for by the correlation of the two variables (r = -.412, p < .001). However, in the gift buying situation the association between dogmatism and innovation explains only 1.1% of the variance, with a

nonsignificant r of .107.

TABLE 3  
2 X 2 FACTORIAL ANOVA TABLE

Source	DF	Mean Square	F	P
Situation	1	2.132	.915	n.s.
Dogmatism	1	.859	.369	n.s.
Situation X Dogmatism	1	28.341	12.163	.001
Residual	106	246.981	-	-
TOTAL	109	278.364		

Discussion

The results show rather dramatically that manipulation of the purchase situation does have a significant impact on the dogmatism-innovation relationship. The basic inverse dogmatism-innovation relationship was once again found in the purchase for self condition. However, in the gift giving situation a direct relationship was found to exist. High dogmatics were found to be significantly more innovative than low dogmatics. Furthermore, even though the low r between dogmatism and innovation given gift buying is not significant, it is positive. This complete change in direction provides strong evidence that the task condition of self vs. gift purchasing does have considerable impact on the dogmatism-innovation relationship.

As stated earlier, the gift giver must make inferences about the recipient and the potential reaction to the gift. He must also be concerned about the information the particular type of gift chosen will convey about the giver-recipient relationship and, indeed, about the gift giver himself. It may be that higher dogmatic people wish to be seen by gift recipients as being more novel, more innovative than they in fact are. Hence, the purchase choice would be more likely to reflect the generally more valued characteristic (though more personally risky) of being a leader and innovator. On the other hand, it may be that they perceive others as being more novel than themselves and, therefore, buy more innovatively for them.

Without more research on the specific aspects of the gift giver situation per se, it is difficult to explain fully why the dogmatism-innovation relationship normally found is reversed. It is highly significant, however, to note that the direction of the relationship did change, as the basic intent of this research effort was to indicate that the situation does play a powerful role in the buying environment. Research must include situational variables, such as those defined by Belk (1975b) if one is to accurately portray real consumer purchase decisions.

References

- Roger G. Barker and Allan W. Wicker, "Commentaries on Belk, Situational Variables," Journal of Consumer Research, 2(December, 1975), 165-168.
- Russel W. Belk, "An Exploratory Assessment of Situational Effects in Buyer Behavior," Journal of Marketing Research, (May, 1974), 156-163.
- \_\_\_\_\_, "Situational Variables and Consumer

Behavior," Journal of Consumer Research, 2(December, 1975a), 157-164.

\_\_\_\_\_, "Reply to Barker and Wicker," Journal of Consumer Research, 2(December, 1975b), 235-236.

\_\_\_\_\_, "It's the Thought That Counts: A Signed Digraph Analysis of Gift Giving," Journal of Consumer Research, 3(December, 1976), 155-162.

Andre Caron and Scott Ward, "Gift Decisions by Kids and Parents," Journal of Advertising Research, 15(August, 1975), 15-20.

Kenneth A. Coney, "Dogmatism and Innovation: A Replication," Journal of Marketing Research, (November, 1972), 453-455.

Martin Fishbein and Icek Ajzen, Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research, (Reading, Mass.: Addison-Wesley Publishing Co., 1975).

Jacob Jacoby, "Personality and Innovation Proneness," Journal of Marketing Research, (May, 1971), 244-247.

Harold H. Kassarian, "Personality and Consumer Behavior: A Review," Journal of Marketing Research, (November, 1971), 409-418.

R. J. Lavidge, "The Cotton Candy Concept: Intro-Individual Variability," in Lee Adler and Irving Crespi, Attitude Research at Sea, (Chicago: American Marketing Association, 1966), 39-50.

S. Ward and T. S. Robertson, "Consumer Behavior Research: Promise and Prospects," in Scott Ward and Thomas S. Robertson, Consumer Behavior: Theoretical Sources, (Englewood Cliffs: Prentice-Hall, 1973), 3-42.

## PERSONALITY: THE LONGEST FAD

Harold H. Kassarijian, University of California, Los Angeles

### Introduction

Having been brought forth by Motivation Research and nurtured by psychological theory, the field of consumer research seems destined to be saddled by personality. Somehow if psychosis and the selection of a spouse, suicide and child rearing can be accounted for by personality variables - so, too, it must be that the selection of Sunkist oranges over Florida oranges, preference for brands of canned peas, and exposure to the National Enquirer or LaVerne and Shirley must be accounted for by personality.

We now have some 300 papers and a dozen review articles and bibliographies in the field ranging from the mediocre and out-of-date to the creative and ingenious. And yet the results have remained remarkably stable over the past two decades - five to 15% percent of the variance of the behavior of the consumer can be accounted for by personality.

As one examines the development and emergence of these studies and these researchers, again a remarkable stability becomes obvious. Personality research is similar to research in most other sub-fields of consumer behavior. Let me try to type cast or stereotype researchers in consumer behavior. In doing so, not only do I plagiarize the diffusion research people, but also myself in a paper written at another time for another purpose (Kassarijian, In Press).

### The Scholar

The first of these stereotypes is the true scholar. Highly educated, widely read, a leading expert in his or her own field, by no means is he narrow in scope of interest or in breadth of knowledge. The scholar is trained and influenced more by the European tradition of broad scholarship and adheres much less to the technological tradition in the United States. When necessary and appropriate; philosophy, science, the arts, humanities and mathematics could be brought into his writing or into her thinking. This type of man is able to see the interrelationships between several fields and can link them together. People like Katona, Lazarsfeld, Bauer and Howard qualify as scholars.

### The Innovator

A second type of researcher is the innovator. Bright, sharp, well read in various disciplines, he may be exceptionally creative or he may be a creative borrower, but would be the first to do a study. The quality of the study itself, may not be the most sophisticated possible but it is innovative. Postman and Bruner's early work on selective perception in psychology, or Janis and Feshbach's original study on use of fear appeals in persuasive communications would be examples. Evan's work on Fords and Chevrolets, some of the very early work in cognitive dissonance, Levy's work on social class, Kuehn's work on orange juice, and early contributions in distraction, attribution theory, and modeling all belong here. Better quality research, advanced technology, sophisticated statistics, and complex experimental designs are left to others. But without him, there is no place for methodologists, the sophisticated statisticians, and the experimental design specialists to ply their trade.

### The Middle Majority

Using the terminology of the diffusion people, the third stereotype of researcher I would like to call the middle majority. The middle majority, in fact consists of two simultaneously co-existing breeds of researchers. The first of these is the simple follower. He is not too well trained. But he is alert to the literature and knows what is new and publishable in the field. Once a topic has become a fad, his time has come. There is nothing very innovative in his or her thinking, nothing particularly brilliant in his design, or his analysis. He is simply a follower. More likely he aims at the second tier and third tier journals and is thrilled to have his empirical research published any place and at any time. He contributes little of major consequence, his work is seldom referenced, and the overall contribution he makes is not the architectural design of the edifice of knowledge but he rather furnishes the little building blocks on the facade. The work is important for the advancement of science, but not as highly rewarded as that of either the innovator or the methodologist.

Co-existing with the simple follower is the methodologist or technocrat. Highly trained, mathematically and statistically sophisticated, this person is the technical expert, the experimental design specialist, the consultant to peers and colleagues on methodology. He is capable of advancing the field rapidly as he brings in new variables, more attributes, interrelationships, and the ability to process the masses of data his computers can handle. Sometimes, but not always, he is the pedantic data cruncher with emphasis on the minutiae and the presentation of a very narrow band of knowledge. To others, this work often tends to appear dull and does not excite more than a few peers. Nevertheless the work is of very high quality and will be published until the very end in the highest quality journals in our field. His major contribution, perhaps, is not the advance of science - although he makes advances, no doubt about it - nor is it necessarily in the development of new sophisticated methodologies - although he may make some - but rather in the development of a concept and in hastening its burial.

Once the multi-variables are intercorrelated and controlled, interest in the topic drops; the simple followers have followed the innovators to a new topic and the pedantic data crunchers for the most part will do likewise.

### The Laggard

Finally, the fourth category of researcher is made up of the laggard and the translator. The translator has managed to find public policy or management implications and has written the Harvard Business Review articles and the chapters in textbooks. He is most important, for without him neither the layman in consumer behavior nor the student would ever understand what is new and at the forefront in research. The laggard is the saddest of all. Long after the innovator has come and gone, after the technician has crunched the interest out of a topic, the laggard appears with a lightweight study, several generations of ideas behind the forefront. His statistics and his methodology is often even less sophisticated than that of the innovator, and every thought he has, has been thought before. His work is rejected by

second and third tier journals and he is and remains frustrated, convinced that journal editors and reviewers are biased against "lesser known academics from lesser known schools." If only articles would be reviewed anonymously and only editors were not prejudiced against his or her school, he too would have a chance. He feels very deeply that not being in the inner clique, his work is not appreciated or noticed, and eventually concludes that he no longer wishes to contribute to the "clutter" in the literature.

With that preamble we can turn to the present set of papers.

#### Fornell and Westbrook

The first is Fornell and Westbrook - an exploratory study on assertiveness, aggressiveness, and consumer complaining behavior. The paper includes a very nice concise review of complaint behavior - a topic accumulating more and more interest in recent years.

The authors have rather creatively taken the frustration-aggression hypothesis of Dollard from Learning Theory and assertiveness training from Clinical Psychology and tied it into complaining behavior. A battery of paper and pencil tests, administered to a group of students, was factor analyzed such that the critical scales of aggressiveness and assertiveness (or better titled "non-assertiveness") each consisted of three items. Outside of a bit of face validity in naming the scales, the key problems of validity and reliability are not addressed in sufficient depth. For example, the authors might have given the same instrument to the same subjects 30 days later to get a measure of reliability. Further, additional measures of construct validity other than those used should not have been that much more difficult.

With regard to my overall theme in this discussion, note that the scales of "aggressiveness" and "assertiveness" correlated with complaining behavior with  $R^2$ 's such that from 8 to 16% of the variance is accounted for. Nevertheless the paper does make an important contribution to the literature on consumer dissatisfaction if not to the personality literature. The discussion of assertiveness, aggression, dissatisfaction and frustration are far more important than the data manipulations. In assigning this work to my earlier categories, it would likely be classed as reflecting the creative borrower, the simple follower or the sophisticated technician (or technocrat).

#### Hirschman and Wallendorf

The Hirschman-Wallendorf paper relies not on psychological theory or personality but turns to sociology as social-connectedness, exposure to the mass media, and cosmopolitanism are interrelated. The inter-correlations range from .08 to .47, the levels of significance from .03 to .001, while the percent of variance accounted for by personality (cosmopolitanism) ranges between 2% and 20%.

Once again, in the tradition of the literature on consumer behavior, the validity and reliability data are not emphasized in the paper (although multi-method multi-trait methodology in the mold of Campbell and Fiske are employed in the design). The authors have had to use some operational (and therefore arbitrary) definitions of role accumulation, of media exposure and of cosmopolitanism. Further, in the tradition of the technocrat, we see factor analyses with factor loadings carried out to five digits beyond the decimal. Theory of significant digits - where are you when we need you?

The major contribution of this paper is not in the data analysis or predictive ability of still another personality measure but rather in the ingenious introduction of sociological concepts into consumer behavior. As to the classification of the researchers, I suggest the audience consider innovator, creative borrower or sophisticated technician as possibilities.

#### Coney and Harmon

The greatest contribution of the Coney and Harmon paper is that it is a replication and extension of the dogmatism and innovation literature. They seem to be following Nakanishi's brilliant recommendations (1972) on personality research.

Nakanishi, in a most insightful paper presented at the 1972 meetings of the Association for Consumer Research, has suggested that the low explanatory power of personality characteristics may have stemmed, in part, from naive conceptualizations of the relationship between personality and consumer behavior often held by researchers in the field. Nakanishi suggests that personality is better conceived as a dynamic concept, not a consistency over a variety of situations but rather consistency in the manner the individual adjusts to change over time and over situations. He writes it is perhaps, "more correct to conceive of personality as a moderator variable whose function is to moderate the effects of environmental change in the individual's behavior. This dynamic concept of personality has not been taken seriously in personality research."

Hence, the relevant variables involve in the least, personality traits, response and behavior patterns, moderator variables, situations and individuals. Furthermore, for some of these variables it is essential that measurements be taken over time. That is, as we sample individuals, traits and responses, we should also take samplings of situations over time. Coney and Harmon find the significant interaction between situations and personality predicted by Nakanishi.

Although the Rokeach dogmatism scale is more or less validated the innovation scale used by the authors is not. The data analysis is by no means such that the authors qualify as technocrats, and in fact the analysis of variance used does not even quite feel right in this design. Nevertheless sophisticated data analysis and computer output will not alter the obviously significant interaction effect or the basic contribution of this paper as a replication and extension. In summary this paper neither qualifies as the work of the innovator nor really that of the laggard. I suggest it is middle majority or maybe late majority.

#### Conclusions

In summary, we see once again that personality, by itself, does not predict behavior all that well and even when treated as a "moderator variable" the results are not all that powerful. The conclusion remains that about ten percent of the variance can be accounted for by personality variables. I do not mean to suggest or imply that predictability is an important, necessary, or sufficient condition in personality research; but rather that we probably need no more personality studies of the naive variety. We have milked and massaged this topic for all it is worth. The next step must not be another test instrument correlated with purchase or another product or related to another moderator variable, but some major studies along the combined cross sectional and time series lines suggested by Nakanishi and suggested by the three papers in this session.

It is here that the technical specialists that abound in our field can make contributions. With sophisticated computer hardware, complex software, and multi-dimensional statistical models, the fine variations and details can be argued, presented and rejoined. As for myself, I am tired of the topic. I am fully prepared to accept that 10% of the variance in the behavior of the consumer can be accounted for by personality. I leave it to others to examine the 10% and demonstrate that more properly it should be 5% or 15%. But, I, for one, hope to turn to the remaining 85 or 95 percent of the variance that personality will never be able to explain.

#### References

Harold H. Kassarian, "Marketing, Consumer Behavior and Raymond A. Bauer," in Proceedings of The Tenth Paul D. Converse Marketing Symposium, Alan Andreasen, ed., (Chicago: American Marketing Association), In Press.

Masao Nakanishi, "Personality and Consumer Behavior: Extensions," in Proceedings of the 1972 Meetings of the Association for Consumer Research, M. Venkatesan, ed., 1972, 61-65.

Jerome B. Kernan, University of Cincinnati

## Introduction

Over the years, many of us have professed a vow to one of consumer research's cardinal articles of faith--that behavior is a (however convoluted) function of personality. Just as "there must be a God," a person's behavior "must" derive, in some meaningful way, from his/her idiosyncratic self. That this relationship has never been proved (that faith is the only basis for accepting it) has not deterred the personality disciples from their evangelical ministry; after all, this is the sort of thing that any right-headed (no reference to Krugman) person knows is correct. Besides, if "faith can move mountains" it surely can withstand the curious barbs of a few unwashed existentialists.

Personality cognoscenti are a clever lot; we know where and how we are culpable. So we repair to the inoculation game (McGuire, 1969; Bither et al., 1971), taking care to place our most credible sources in the vanguard (Wells, 1966; Jacoby, 1969; Kassarian, 1971). The result, of course, is thoroughly predictable; we are thereafter left free to purvey our prestidigitation with impunity. Indeed, we are encouraged; the American Marketing Association--hardly known for its support of unpopular causes--has bestowed the ultimate imprimatur on us by issuing a bibliography on "personality research" (Twedt et al., 1977). The whole thing at times seems a bit like sinners defining evil.

But all this is innocent enough, and infinitely more interesting than teaching impatient students, running dreary crosstabs for persistent clients, or getting hassled by your section chief for deviating from OMB's latest procedural directive. And occasionally we even learn something.

So what have we here? The Hirshman/Wallendorf paper is not in the "Sociological Perspectives" session: Coney and Harmon are not in the "Situational Impacts...Gift Giving" session; and "Consumer Satisfaction/Dissatisfaction" must survive without Fornell and Westbrook. Three displaced papers? Perhaps; but only if one focuses on their object sets. If we regard these papers contextually, they are disparate--adoption/diffusion, gift-giving, and complaining would hardly load on the same factor (unless it's "consumer behavior"). But if we look carefully, a common theme emerges; each of the papers essays a trait/behavior connection. Not each study designates its trait(s) as "personality" and the measures of behavior are not uniformly explicit. But at least by implication each of the researches proffers a trait → behavior relationship; each falls within the psychology-of-individual-differences domain.

Following are some comments to which a variety of valences can be attached. They are benign items--just matters potentially worth consideration.

## Coney/Harmon

There are two particularly encouraging things about this study. First, it advances the cause of replication, something sorely needed in our empirical work. And second, it emphasizes the importance of environmental circumstances on consumer behavior, demonstrating the obvious (but so often ignored) fact that people are not trans-situational automatons. Kissing your lover is not the same as kissing your sibling.

Some issues about which one might become a bit anxious:

- The experimental subjects--student volunteers--and whether they behaved naturally.
- The criterion measure (purchase intention) vis-a-vis purchase, per se; Fishbein notwithstanding, everybody realizes the two can be very different depending on (what else?) the equivalence of the situations under which each is measured.
- The innovation measure; this is a mystery unless one happens to be familiar with the relevant Jacoby (1971) and Coney (1972) studies. Presumably, the chance-alone mean value is 3.0--12 products x .25.
- Whether the most appropriate ANOVA model has been used; an argument might be made that dogmatism is nested within subjects while situations are not, hence the design is a mixed one and not the factorial as used.

## Fornell/Westbrook

This study is interesting for its attempt to carry an aspect of personality theory through a counseling-psychology paradigm to a consumer (complaining) behavior context. Whether it produces dramatic results is not so important as the fact that it is attempted; we need--and are likely to see--more such behavior-is-a-process work in the future. The study represents an interesting conceptualization, one that is likely to stimulate replications and extensions. And what might those researches attempt to establish? Two possibilities would seem obvious:

- Whether submissiveness/assertiveness/aggressiveness are distinct traits or simply differences in degree of a trait (presumably one born of frustration or anxiety) and, depending on that answer, whether effective complaining (measured non-hypothetically) is a non-monotonic function of such a continuum or of something(s) altogether different. Equity theory (Berkowitz and Walster, 1976) would seem to offer some useful insights here.
- What "changes" when a non-student sample is used and assertiveness and complaining are not measured in the same setting? Granting all the reasons for using a student sample we all realize that this practice imposes severe limitations. Students "...chosen selectively to represent all class levels and both sexes" is virtually a contradiction in terms; all class levels simply can't be found at "...a major private university" and, even if they could, students are notoriously better connected to the social/industrial/political structure, have measurably superior communication skills and proclivities and, in general, are just more "together" than "plain folks." A priori, students can be expected to be more assertive (it's a cognitively complex activity) than a consumer population in general (an alarming proportion of whom are functionally illiterate), and that fact affects both absolute and relative associations (student scores have much less elevation).

This research, although approaching the matter from a slightly different perspective, is reminiscent of the Reynolds and Darden (1971) model, which also argued that the more socially-connected a person is, the more s(he) is likely to behave in innovative ways. Now, as in 1971, that contention seems a bit awkward to dispute. Let us therefore accept it and comment briefly on some methodological points. The ultimate implications of this work--role strain isn't so intolerable (Sieber, 1974) and societal organizations pose threats to gradual social change (Blau, 1974)--are so value-laden that to discuss them is to add yet another role, polemicist.

The methodological comments, if a guess is permissible, all hang on the same factor--reanalysis of already-collected data. Thus one can appreciate that certain constraints existed; yet some legitimate concerns remain:

- Is role accumulation per se the variable of interest or should we be concerned with that constraint as it is a perceived instrumentality? Which is more important--that one joins groups or why one joins?
- Can contamination occur in operationalizing role accumulation? Suppose I belong to a country club. Given that a choice between recreational and social organization can be made, do I "get credit for" golfing, tennis, bridge, and spectator-sports events (I do each, but with the same people)? Isn't an alternative (heaven forbid, we've been brainwashed by expectancy-value models) to regard organizational membership as a degree of institutionalization for an activity?
- Several curious judgments seem to be reflected in the operationalization of the variables. For example, in what sense is the "reduced set" of TV program types more "stimulation rich" than the original set; what happened to "variety" as a criterion for stimulation? Do demographic and socioeconomic variables count for naught; are the measured variables mediated equivalently regardless of who a person is?

#### Directions For Future Research

At this juncture it is a capital offense to reiterate the do-and-don't litany of personality research in consumer behavior. Every graduate student has read and/or heard it, ad nauseum. No harangues, then; just four modest proposals.

#### Beware of Operationalized Constructs

We must be realistic. Personality is a construct and, as such, has no real-world referent. Accordingly, it can be operationalized only according to someone's invention. Unlike the concept, age (which can be measured with both trans-situational and trans-personal consensus), personality "means" something different according to who measures it and how. As a predictor variable, then, "personality" is of questionable validity and reliability. And we don't reify it by multi-variately laundering a set of existing traits or ad hoc personality-like statements. Molar sets of traits "work better" than individual traits, but unless the predictor is "inclination to buy Brand X," one should not expect much of Brand X behavior to be explained by the predictor.

Typically, personality/behavior associations are weak. That is, the measures imply weak associations. But what does that tell us? Of course, the association might indeed be weak. But often as not the "weakness" proceeds from unstable or irrelevant criterion variables (how much confidence can we place in self-report measures?), invalid or unreliable predictor variables (as noted above), a non-linear relationship, or an indirect one. We know about each of these possibilities, yet seem oblivious to them when the quick-and-dirty-study occasion presents itself. Would any of us really believe a high  $r^2$ ?

#### Who's Kidding Whom?

Unless one is a bona fide personality researcher (i.e., personality qua personality) and/or wishes to opt for the sanctuary of the ivory tower (where "exploratory" is code for "anything goes"), it's patently obvious that personality research is a means to an end--personality mediates, but doesn't cause behavior. Accordingly, personality should be assessed in loco, within the context its effect is suspect. This means testing respondents like those whose behavior is under inquiry and in real situations. (It also implies that the typical student-sample study, replete with its demand characteristics, is fatuous.)

#### If You Want to Play Hardball. . .

Beyond the world of hassling with journal editors there is a land in which the name of the game is predictive validity. In that world, personality either works (meaning it has an instrumental value in identifying behavioral patterns) or it is discarded (it's true relationship to the behavior in question is so camouflaged as to be useless). Although the success rate varies widely, the research paradigm seems to be fairly well established.

- First, the policy issue or administrative problem is identified and some prospective renditions of criterion variables as well as research hypotheses are specified.

- Next, a series of focus-group or depth interviews is conducted to (1) ascertain the validity of the notions originally specified and (2) translate the phenomenon into respondent vernacular. The interviewers must be qualified and the participants representative of whatever population whose behavior ultimately is to be explained. For instance, one might inquire at this stage: "What sorts of problems do you encounter with (product)? What do you do when you feel you've been treated unfairly? Whose fault is it when a product doesn't work? Etc."

- Next, a set of variables--some of which are personality-like--is gleaned from the previous interviews and pilot-tested, in a structured format, on a group of perhaps 100-200 people who reflect all the important characteristics of the target population. It should be emphasized that situational, demographic, socioeconomic and product (service)-specific variables are tested simultaneously at this stage.

- Another set of qualitative interviews might then be necessary to verify any surprises uncovered in the pilot survey.

- The results are finally used to construct a standardized questionnaire for a statistical survey



(n = 1000-1500). The output usually includes a small portion directed to the creative group and it says something like: "Given these palpable measures of the various population (market) segments, the following personality data suggest which appeals will be most effective for each segment."

#### Is It Worth It?

There is increasing skepticism being expressed about personality research in consumer contexts. Even psychographics is questioned. This is all a matter of personal opinion, of course, but one view is that there is more to the world than r's and that the study of personality can be evaluated apart from a criterion of prediction. Another way to pose that is to ask yourself whether sex that does not result in conception is functional.

#### References

- Leonard Berkowitz and Elaine Walster, eds. Equity Theory: Toward a General Theory of Social Interaction (New York: Academic Press, 1976).
- Stewart W. Bither, Ira J. Dolich and Elaine B. Nell, "The Application of Attitude Immunization Techniques in Marketing," Journal of Marketing Research, 8 (February 1971), 56-61.
- Peter M. Blau, "Presidential Address: Parameters of Social Structure," American Sociological Review, 39 (October 1974), 615-635.
- Kenneth A. Coney, "Dogmatism and Innovation: A Replication," Journal of Marketing Research, 9 (November 1972), 453-455.
- Jacob Jacoby, "Personality and Consumer Behavior: How Not to Find Relationships," Purdue Papers in Consumer Psychology, No. 102, 1969.
- \_\_\_\_\_, "Personality and Innovation Proneness," Journal of Marketing Research, 8 (May 1971), 244-247.
- Harold H. Kassarian, "Personality and Consumer Behavior: A Review," Journal of Marketing Research, 8 (November 1971), 409-418.
- William J. McGuire, "The Nature of Attitude and Attitude Change," in Handbook of Social Psychology, Vol. 3, Gardner Lindzey and Elliot Aronson, eds. (Reading, Mass.: Addison-Wesley, 1969), 136-314.
- Fred D. Reynolds and William R. Darden, "Mutually Adaptive Effects of Interpersonal Communication," Journal of Marketing Research, 8 (November 1971), 449-454.
- Sam D. Sieber, "Toward A Theory of Role Accumulation," American Sociological Review, 39 (August 1974), 567-578.
- Dick Warren Twedt, Lyndon E. Dawson, Jr., Hugh G. Wales, and Gary N. Bunner, eds., Personality Research in Marketing: A Bibliography (Chicago: American Marketing Association, 1977).
- William D. Wells, "General Personality Tests and Consumer Behavior, in On Knowing the Consumer, Joseph Newman, ed., (New York: John Wiley & Sons, Inc., 1966), 187-189.

FAMILY LIFE CYCLE AS A DETERMINANT OF SIZE  
AND COMPOSITION OF HOUSEHOLD EXPENDITURES

Johan Arndt, Norwegian School of Economics and Business Administration

Abstract

This study is concerned with the family life cycle as a predictor of spending patterns. In a national probability sample of Norwegian households stage in family life cycle was related both to size of expenditures and their allocation to product and service groups and to gifts given.

Introduction

In consumer behavior the family life cycle notion is one of the most over-quoted and under-researched concepts. The family life cycle is an idealized construct referring to the fact that the typical family passes through a fixed sequence of important stages in its life. These stages affect the economic behavior of the family.

Most consumer behavior textbooks, such as Engel, Kollat, and Blackwell (1973, pp. 193-196), Block and Roering (1976, pp. 138-142), and Berkman and Gilson (1978, pp. 204-206), routinely discuss the life cycle concept and its implications. But the studies quoted tend to be dated ever-greens. In fact, some of the most important contributions on the subject matter were presented at a conference at Ann Arbor, Michigan a quarter of a century ago (Clark, 1955). For instance, in one major paper, Lansing and Morgan (1955) showed how income, expenditures on durable goods, assets and debts, and subjective feelings about the financial situation depended on stage in the life cycle. Other relevant papers were delivered by Barton (1955), Fisher (1955), and Miller (1955). These and other studies such as Lansing and Kish (1957) were reviewed a decade ago by Wells and Gubar (1966).

However, after this promising start, there has been a long hiatus in which few memorable new studies have emerged. An exception is the mainly unsuccessful attempt by Rich & Jain (1968) to explain variations in women's shopping by stage in life cycle. A partial explanation of the neglect of the family life cycle is the dominating interest in studies having the individual consumer (rather than the household) as the unit of analysis. Another reason may be the shift away from single constructs such as the family life cycle to more comprehensive theories.

The purpose of this article is to contribute to a recycling of the family life cycle concept. As appears from the Wells and Gubar (1966) review, past contributions have mainly treated the life cycle as an independent variable to account for differences in stocks (product ownership and debts) and flows (incomes and expenditures of different kinds).

Past studies have mostly been concerned with isolated, fragmented aspects of economic behavior. This study attempts to make a more "holistic" approach in the sense that it addresses the total amount of household expenditures and its allocation to products and services bought, taxes, social insurances, and fees paid, and gifts given. A limitation of this study (and its predecessors) is that it is essentially a distributive study focusing on behavioral outcomes of decision processes rather than the process leading to the outcomes (Lazarsfeld, 1959).

Before describing the method and presenting the results

of the study, a few comments on the theoretical importance of the family life cycle concept are in order.

Family Life Cycle And Spending Outcomes

The family life cycle notion refers to the important transitions or critical passages in the life of the typical nuclear family. Inside the family, the social roles of its members differ over the life cycle. Moreover, externally, the social position of the family in relation to society is partially determined by stage in the life cycle.

In the pre-industrial society, the family (mainly the extended family) was not only the appropriate unit of analysis for consumption activities, also most of the production functions were carried out in the context of the family. In modern society these two functions have been split, leaving only consumption activities as the domain of households or families. Even in this area there are indications that the role of the household is decreasing (dining out, prepared foods, etc.).

To the student of consumer behavior, family life cycle is of some interest as it is related both to the economic resources (purchasing power) of the family and to the needs and wants manifested in the spending of the money.

Lansing and Kish (1957) and Lansing and Morgan (1955) found that over the life cycle the total incomes of the spending units increased to a peak and fell off in the last stages. However, there was a definite dip for young married couples with young children due to the withdrawal of wives from the labor force.

Family life cycle affects the composition of consumer expenditures in two ways. First, there is the income effect referring to the fact that products and services differ in income elasticity (as pointed out, incomes vary over the life cycle). Necessities such as food, clothes, and minimum housing expenses are fairly inelastic with regard to income. However, income increases result in disproportionate increases for "luxury" or "discretionary" items such as cars, expensive appliances, recreation items and gifts given. Second, stage in family life cycle may be used as a proxy or surrogate indicator of needs and wants. For instance, young couples with children have a higher need for food and clothing, and may be expected to spend a disproportionate part of their incomes on these categories.

The simplified framework underlying the discussion above is summarized in Table 1.

A recurring problem with family life cycle research is the difficulty of developing a meaningful categorization of the life cycle. As observed by Wells and Gubar (1966, p. 360), not two investigators have yet agreed on which categories to use. Needless to say, this has hindered comparisons and the development of a cumulative research tradition. In brief, the trade-off problem is as follows: There will be too large within-class variation (and too little between-class variation) if the taxonomy contains a small number of classes. On the other hand, a large number of categories may mean that many classes

may be unpopulated unless large-scale surveys are involved.

Table 1

SIMPLIFIED CONCEPTUAL FRAMEWORK OF THE RELATIONSHIP BETWEEN FAMILY LIFE CYCLE AND EXPENDITURES

	<u>Antecedent Variables</u>	<u>Consequent Variables</u>
Concepts	Economic resources Needs and Wants	Household income Household demand
Indicators	Stage in family life cycle	Size and composition of household expenditures

In the study to be presented here, the following six-class scheme was used: (1) Bachelor, (2) Newly married, (3) Full nest I, (4) Full nest II, (5) Empty nest, and (6) Soletary survivors. The criteria are shown in Table 2 below. The categorization which partly was determined by the secondary nature of the data base, differs somewhat from the nine-category scheme recommended by Wells and Gubar (1966). The differences are as follows: Wells and Gubar (1966) recommend distinguishing (1) between Full nest I families with youngest child under six and others, (2) between Empty nest families with head of household in labor force and others, and (3) between Soletary survivors in labor force and those who are retired.

Table 2 below summarizes the theoretical expectations for this study based on the general discussion above, the specific findings reviewed by Wells and Gubar (1966), and some common-sense. The central idea is that economic resources differ by stage in family life cycle, and that each stage has an appropriate "standard package" of goods and services.

Method

Ideally, a longitudinal (panel) study would be the appropriate research strategy for examining consumer behavior over the family life cycle. Such an approach would make it possible to observe changes in income and demand composition when they occur. Unfortunately, such an approach is infeasible in practice because of the high costs involved, and, more important, the long time perspective required - literally the whole professional life of a researcher.

Hence, this study like preceding studies (Wells and Gubar 1966) had to resort to a less than ideal cross-sectional design. Such an approach permits making inferences about what happens at the various life cycle stages by comparing households in different stages.

The source of the data was the Survey of Consumer Expenditure 1973 conducted by the Central Bureau of Statistics of Norway. An original purpose of this study was to update the weights used to calculate the consumer price index. The unit of analysis in this survey was the household, defined as all persons having a common

TABLE 2

ECONOMIC RESOURCE SITUATION AND COMPOSITION OF CONSUMPTION EXPENDITURES BY HOUSEHOLDS IN DIFFERENT STAGES OF THE LIFE CYCLE

Stage of Family Life Cycle	Criteria	Resource Situation	Composition of Consumption Expenditures	Gift-Giving
Bachelor	Young (less than 40) single.	Low income, one-person households, early stages of career.	Overspend on restaurants, cars, and recreation. Spend little on food, housing, and furniture and appliances.	Fairly high amount allocated to gifts because of high discretionary income.
Newly Married	Young (main income earner less than 40) couples without children	High income. Two income earners.	Spend more than Bachelors on food, housing, and furniture and appliances, but less on restaurants.	Spend relatively less than Bachelors, in spite of higher income. Need money to establish the household.
Full Nest I	Young (main income earner less than 40) couples with children (less than 16).	Lower income than former group because more likely to have one income earner.	Overspend on food and housing while spending less on cars and restaurants.	Like former group.
Full Nest II	Older (main income earner 40 or more) couples with children (less than 16).	Higher income than former group. Members of this group are more likely to be further in career and having two incomes.	Overspend on food, furniture and appliances (second generation), and recreation.	Like the former two groups.
Empty Nest	Older (main income earner 40 or more) couples without dependent children (less than 16).	High income.	Overspend on recreation, beverages and tobacco, furniture, and appliances.	High degree of gift-giving.
Soletary Survivor	Older (40 or more) single.	Relatively low income, High proportion of retired and also low education persons.	Overspend on food, housing and medical care. Spend little on cars, recreation, and restaurants.	High degree of gift-giving because most of own needs are met.

dwelling and sharing at least one meal per day, excluding all institutional households such as hospitals.

#### Sample

The sample consisted of 4,707 households selected by a stratified, self-weighting two-stage probability sampling method, ensuring a nationally representative sample of households.

Each of the households living at the 4,707 drawn addresses were contacted by interviewers and asked to participate in a personal interview (collecting information about name of household members, date and year of birth, household and marital status, and occupation and occupational status). The second source of data was detailed account books in which respondents (in return for a small honorarium) were requested to enter all expenditures over a two-week period on products and services bought, payments for taxes, social insurances, and fees, and the value of gifts given. Consumption of own-produced products (milk, butter, potatoes, fresh meat, etc.) were also recorded. Third, when the account books were collected, a second personal interview gathered information about expenditures on products with low purchase frequencies (including certain expenses for dwelling, car purchase, and purchase of special appliances) and information about product ownership. As size and composition of household expenditures vary with season, data had to be collected over a whole year. For instance, 1/26 of the households kept accounts between January 1 and 14, 1973, another 1/26 between January 15 and 28 and so on.

In all, 3,363 households (or 71 per cent) completed the account books and participated in both interviews. The most important reasons for non-response were refusals to participate and absence from the address.

#### Measures

Stage in Family Life Cycle. The 3,363 households were classified by stage in the life cycle on the basis of the demographic information from the introductory interview. The distribution is shown in Table 3. A reason for the large relative size of the "Other Households" class was the fact that the upper age limit for dependent children was 16 years (set by the Central Bureau of Statistics).

Table 3

PROPORTION OF HOUSEHOLDS IN DIFFERENT STAGES OF FAMILY LIFE CYCLE

Stage	Per Cent
Bachelor	2.8%
Newly Married	3.1
Full Nest I	19.5
Full Nest II	9.1
Empty Nest	17.9
Solitary Survivor	13.4
Other Households	34.2
Total	100.0%
Base	3,363

#### Expenditures

The account books were the primary source of expenditure data. The record-keeping was very detailed. For instance, distinction was made between five different

varieties of bread. Here we will be concerned with (aggregated) main groups of consumption expenditures (CE). It was here distinguished among the following nine expenditure groups (EG): (1) Food, (2) Beverages and tobacco, (3) Clothing and footwear, (4) Housing (rent, fuel, and power), (5) Furniture and household equipment (appliances), (6) Medical care, (7) Transport (car, gasoline, public transportation, etc.), (8) Recreation and education (television sets, boats, movies, theaters, books, etc.), and (9) Other goods and services (restaurants, hotels, packaged tours, personal care articles, etc.). For each household, the annual expenditures in each of the nine groups were projected.

Second, all expenditures on taxes, social insurances, and fees (TSF) were also recorded in the account books, along with the monetary amount of gifts given (G).

The total household expenditures (TE) were accordingly composed as follows:

$$TE = CE + TSF + G \quad (1)$$

where

$$CE = \sum_{i=1}^9 EG_i$$

In this study total household expenditures (TE) may be treated as an indicator of household income. It should be noted, however, that savings and investments in real estate are not included in TE. However, TE is believed to be strongly correlated with household incomes.

#### Results

The size and composition of consumption expenditures by stage in family life cycle are shown in Table 4. As seen in Table 4, the average consumption expenditures (CE) per household were Norwegian Kroner 36,832 in 1973 (a Norwegian Kroner is about U.S. \$20). Compared with the findings from the corresponding survey in 1967, this indicates an increase in expenditures of 75 per cent. During this period the consumer price index rose by 44 per cent.

The most important expenditure group was Food having a budget share of 24 per cent. This is a substantial relative decrease from 1967 when the food share was 30 per cent. On the other hand, the shares of Recreation and education (9.7 per cent) and Transport (19.5 per cent) increased significantly from 1967. Such patterns as found here are common in societies in fairly rapid economic growth.

Below, first the results relating to size of expenditures will be discussed. Next, findings with regard to the composition of consumption expenditures and gift-giving will be presented.

#### Size of Expenditures

Total household expenditures (TE) are shown in Table 5 (4th row). Like in the earlier American studies (Lansing and Morgan, 1955 and Lansing and Kish, 1957), the lowest household incomes were at the extreme ends of the scale, among Bachelors and Solitary survivors. The relative levels of the lowest to the highest incomes were also similar to the U.S. patterns. Moreover, in support of the theoretical expectations (and earlier empirical findings), there was a dip in the expenditures curve for the Full nest I stage. The highest total expenditures were found for Newly marrieds. Also Full nest II families had high expenditures.

Table 4

RELATIONSHIP BETWEEN FAMILY LIFE CYCLE AND SIZE AND  
COMPOSITION OF CONSUMPTION EXPENDITURES IN 1973

Expenditure Group	Stage In Family Life Cycle							TOTAL
	Bachelor	Newly Married	Full Nest I	Full Nest II	Empty Nest	Solitary Survivor	Other Households	
Food	13.6%	15.5%	23.1%	24.9%	25.6%	26.3%	24.0%	24.0%
Beverages and Tobacco	5.3	5.4	4.4	4.2	5.8	4.3	5.1	4.9
Clothing and Footwear	10.9	7.3	9.7	11.7	9.2	10.3	11.7	10.5
Housing (rent, fuel and power)	12.9	16.3	14.0	13.1	12.9	20.7	8.8	12.9
Furniture and Household Equipment	7.1	12.9	10.6	9.1	9.5	10.4	8.1	9.4
Medical Care	1.4	1.0	1.9	2.1	3.0	3.5	1.9	2.3
Transport	24.8	24.5	21.3	17.9	18.5	9.1	22.6	19.5
Recreation and Education	11.5	8.6	9.6	10.0	9.4	9.2	10.1	9.7
Other Goods and Services (Restaurants, cafes, hotels, packaged tours)	12.5	8.3	5.4	7.0	6.1	6.1	7.6	6.8
Total	100.0%	99.8%	100.0%	100.0%	100.0%	99.9%	99.9%	100.0%
Base: Total Consumption Expenditures (CE) in Norwegian Kroner	22,636	44,779	44,786	46,621	29,160	15,247	42,591	36,832
Number of Households	95	104	656	307	601	450	1,150	3,363

Table 5

RELATIONSHIP BETWEEN FAMILY LIFE CYCLE AND SIZE AND  
COMPOSITION OF TOTAL HOUSEHOLD EXPENDITURES IN 1973

Expenditure Group	Stage in Family Life Cycle							TOTAL
	Bachelor	Newly Married	Full Nest I	Full Nest II	Empty Nest	Solitary Survivor	Other Households	
Total Consumption Expenditures (CE) - from Table 4	22,636	44,779	44,786	46,621	29,160	15,247	42,591	36,832
Taxes, Social Insurances, and Fees (TSF)	10,281	20,380	17,178	18,356	9,021	4,569	13,301	12,719
Gifts Given	1,290	1,276	1,158	1,354	1,543	1,174	1,739	1,453
Total - Total Household Expenditures (CE + TSF + G)	34,207	66,435	63,122	66,331	39,724	20,990	57,631	51,004
Gifts as Percentage of Total Household Expenditures (G/TE)	3.8%	1.9%	1.8%	2.0%	3.9%	5.6%	3.0%	2.8%
Number of Households	95	104	656	307	601	450	1,150	3,363

As seen in Table 5 (second row), taxes, social insurances, and fees (TSF) made up about one-quarter of total household expenditures. But this proportion varied with stage. Hence, in spite of having 5 per cent lower total expenditures (TE) than Newly marrieds, Full nest I households had approximately the same consumption expenditures (CE) because of lower taxes. However, viewed differently, since there are more mouths to feed in this stage there were, of course, significantly smaller consumption expenditures per family member. Hence, in spite of the tax

advantage for this group, the standard of living may have been lower than in the previous stage.

## Composition of Consumption Expenditures

As predicted, Bachelors were found to spend disproportionately much on "luxuries" such as restaurants and entertainment (Other goods and services in Table 4), cars (Transport), and Recreation and education, while relatively little on Food, Housing, and Furniture and House-

hold equipment.

In contrast, Newly marrieds though having almost double the income spent relatively more on Food, Housing, and Furniture and household equipment. On the other hand, they saved relatively on Clothing and footwear, Recreation and education, and Other goods and services.

When the first child arrives and the household is moved on to the Full nest I stage, the result was as expected higher shares of consumption expenditures for Food, Clothing and footwear, and Housing compensated by savings on Transport, and particularly Other goods and services (apparently curtailing eating out substantially).

The results conform with the expectations also for the fourth stage. Full nest II households reported spending relatively much on Food and Clothing and footwear. However, the predicted high share for Furniture and household equipment did not show up.

The most dramatic change from stage four to five related to total expenditures which shrunk by about 40 per cent. As seen in Table 4, the Empty nest group exhibited many of the patterns predicted for the last stage: high shares for Food and Medical care, and low shares for Recreation and education and Other goods and services. Another surprising finding was the high share for Beverages and tobacco. This group contained a fairly high share of economically inactive persons (households) and also persons with relatively low education.

For the Solitary survivors group the results were as predicted. This group had the lowest household income combined with the highest share of Housing costs, what suggests severe problems in making ends meet.

#### Gift-Giving

The findings relating to gift-giving are shown in Table 5 (third and fifth rows). It appears that in absolute terms, expenditures on gifts were surprisingly constant over the life cycle. However, taken as a proportion of total expenditures it appears that the Newly marrieds and the Full nest I and II households as expected had least to spare. The by far most generous group were Solitary survivors, who, as earlier pointed out, seemed to be in the least favorable economic situation. This may be a parallel to the biblical widow who was generous in spite of her poverty. Another explanation is that this group may also contain persons who feel that most of their own needs are met and, therefore, have surplus money.

#### Concluding Comments

In general, the results of the study tended to support the notion that size and composition of household expenditures are systematically related to stage in family life cycle.

The findings of this and earlier studies have obvious implications for welfare programs and consumer policies. For instance, public consumer programs are likely to have different impact depending on stage in life cycle of the citizens involved. Since consumer resources and needs vary over the life cycle, consumer information programs should be tailored to fit these differences.

Marketing managers may use family life cycle as one out of several inputs in forecasts. Demand for different product and service categories may be estimated from knowledge of the relationship between demand and stage in life cycle and the predicted number of households in the various stages.

Second, family life cycle may be used as a criterion for segmenting markets, identifying the most likely buyers.

The nuclear family is presently under pressure. In the future, it may be increasingly difficult to classify consumers by stage in family life cycle because of informal cohabitation arrangements, and a growing number of divorces and remarriages. For the time being, however, the concept seems to be meaningful in consumer research.

#### References

S. G. Barton, "The Life Cycle and Buying Patterns," in Consumer Behavior, Vol. 11. The Life Cycle and Consumer Behavior, ed. Lincoln H. Clark, New York: New York University Press, 1955, 53-57.

Carl E. Block and Kenneth J. Roering, Essentials of Consumer Behavior, Hinsdale, Illinois: The Dryden Press 1976.

Harold W. Berkman and Christopher C. Gilson, Consumer Behavior, (Encino, California: Dickenson Publishing Company, 1978).

Lincoln H. Clark, ed. Consumer Behavior, Vol. 11. The Life Cycle and Consumer Behavior, (New York: New York University Press, 1955).

James F. Engel, David T. Kollat, and Roger D. Blackwell, Consumer Behavior, 2nd Edition, (New York: Holt, Rinehart and Winston, 1973).

Janet A. Fisher, "Family Life Cycle Analysis in Research on Consumer Behavior," in Consumer Behavior, Vol. 11. The Life Cycle and Consumer Behavior, ed. Lincoln H. Clark, (New York: New York University Press, 1955), 28-35.

John B. Lansing and Leslie Kish, "Family Life Cycle as an Independent Variable," American Sociological Review, 22(1957), 512-9.

John B. Lansing and James N. Morgan, "Consumer Finances Over The Life Cycle," in Consumer Behavior, Vol. 11, The Life Cycle and Consumer Behavior, ed. Lincoln H. Clark, (New York: New York University Press, 1955) 36-51.

Paul F. Lazarsfeld, "Sociological Reflections on Business: Consumers and Managers," in Social Science Research on Business: Product and Potential, eds. Robert A. Dahl, Mason Haire, and Paul F. Lazarsfeld, (New York: Columbia University Press, 1959) 99-155.

Donald L. Miller, "The Life Cycle and the Impact of Advertising," in Consumer Behavior, Vol. 11. The Life Cycle and Consumer Behavior, ed. Lincoln H. Clark, (New York: New York University Press 1975) 61-5.

Stuart U. Rich and Subhash C. Jain, "Social Class and Life Cycle as Predictors of Shopping Behavior," Journal of Marketing Research, 5(1968) 41-9.

William D. Wells and George Gubar, "Life Cycle Concept in Marketing Research," Journal of Marketing Research, 3(1966) 355-63.

## FAMILY LIFE CYCLE AND LEISURE BEHAVIOR RESEARCH

E. Laird Landon, Jr., University of Houston  
William B. Locander, University of Houston

### Abstract

The present research investigates the usefulness of Family Life Cycle (FLC) in the study of leisure/recreation behavior. The survey results presented show that FLC groupings capture much of the variance in recreation behavior. Implications are drawn for both public and private sector decision-makers interested in leisure behavior.

### Introduction

In recent years there has been a growing interest in recreation and leisure research by academics, executives, and public officials (Wells & Gubar, 1966; Omura and Talarzyk, 1975; Voss & Blackwell, 1975). Most of the research to date has focused on describing recreation behavior through large scale surveys. Many of the studies have been atheoretical in that they employ little behavioral theory in the research design.

The present article examines the Family Life Cycle (FLC) concept as a useful tool in understanding leisure and recreation behavior. FLC appears to have much potential for explaining leisure behavior because it matches needs with groups of peoples. That is, FLC offers a construct that is both multidimensional and dynamic. Its multidimensional nature is attributed to the fact that FLC is a composite of several important demographic variables. FLC is dynamic because it accounts for the changing family needs and structure over time. If FLC does capture differences in recreational needs, it would be an extremely useful variable for both recreation product companies and public recreation administrators to use in segmenting markets and predicting demand (Engel, Kollat, and Blackwell, 1978, p. 163-5).

The purpose of this paper is to examine recreation behavior from an FLC perspective. This research attempts to answer four questions:

1. Does FLC significantly relate to the frequency and kinds of leisure behavior?
2. Does FLC help explain a community's usage pattern for public facilities?
3. Can FLC be used to measure perceptions of how adequate are a community's facilities?
4. Will the FLC concept provide a means of helping both public and private sector decision makers analyze leisure markets?

The answers to the above questions might well reveal that researchers should devote more study to FLC as an important correlate of recreation and leisure behavior.

### Previous Research

Some consumer researchers (Reynolds & Wells, 1977) have become interested in leisure research because of the important implications on an individual's use of time, amusement activities, and recreation equipment purchases. Wells & Gubar (1966) indirectly discussed the relationship between leisure/recreation variables and FLC. Family Life Cycle was reported to be better than age for predicting vacation travel and lodging out of

the home city. FLC was correlated with ownership of many recreational objects like toys, sleds, bicycles, boats and trailers. Wells & Gubar quoted literature supporting the notion that for working class men their recreation orientation slowly subsides under the pressure of family obligations and does not return until most of the nest-building demands have been met. Thus, the study of FLC has examined some leisure pursuits, but there has been no work to date examining family life cycle as an exploratory variable relating group needs to recreation activities.

### Method

Interviews in 650 households in Pima County were conducted as part of a larger recreation study for the City of Tucson and Pima County recreation departments. The interviews averaged 45 minutes in length questioning respondents on leisure behavior away from home and about attitudes of existing and proposed facilities. The study was part of a program to develop a master recreation plan for the city and county.

### Sample

The sample was a three stage probability sample. Fifty sampling units were allocated to each of 13 geographical stata, 10 blocks were selected from each statum proportional to size, and 5 households were selected from each block in a systematic (interval) way.

### Family Life Cycle

Table 1 shows the categories suggested by Wells & Gubar. As they indicate there is no consensus among researchers as to the definition of each category. However, in marketing the Wells & Gubar approach has been cited for the longest time. Engel, Kollat, & Blackwell (1978) reproduce the categories in their treatment of FLC.

Table 1 also presents the operational definitions for FLC used in the present study. Several deviations and clarifications should be noted:

1. Researchers have split the age variable at 35, 40, and 45. Forty was used in this study because it is closest to the median household age in the sample.
2. In stages 3 and 4, the young and married descriptions were not operationalized. One could be old and/or single and still be in these stages. Having young children was felt to be more important than age or even marital status. For example, those who were divorced or who are over 40 with children could not be classified if age and marital status were included. The problem with operationalizing stages 3, 4, & 5 as originally described was that they included 10%, 10%, and 15% single people, respectively. To exclude these people from the child rearing stages was thought to be a misclassification. The higher incidence of divorce today was not a theoretical consideration when the original life cycle was described.
3. Stages 8 and 9 were combined because there were few people in each stage. The combined stages improved the appropriateness of Chi-Square tests. However, for leisure activities one might expect these groups to differ since retirement provides time for additional

recreation.

4. Finally, age of children was inferred from a question asking if they were preschool, grade school, or high school age. Even though some ages misclassifications probably resulted, school and grade status are probably more predictive than age. For example, a five and a six year old both in the first grade are probably more alike than two six year olds, one in school and the other not yet in school.

Analysis

FLC was cross tabulated with all demographic variables, recreation activity variables, and facility use and attitude variables. Chi-Square was computed using alpha equal to .05.

Results

FLC was related to nearly all the demographic measures in this study. Of course, many of the relationships are self evident in that variables like occupation, marital status, number of members in household, number of children in different school grades, and age of household head are part of FLC. Other demographics were related to FLC because of their obvious relationship to the FLC defining variables: employment status of other household members, number of years lived in Pima County (related to age), and sex of respondent (married household heads are more likely to have a woman at home). Finally, some demographic variables relate to FLC for conceptual reasons. Occupation, accounting for retirement, was related to FLC. For example, stage 1 was 35% student compared with 4% in the total sample. Owning or

renting was related to FLC in that 28% of stage 1 owned whereas 78% of the total sample did. Likewise, 17% of stage 6 (working over 40, married with no kids) had an advanced degree, whereas ~~34~~ 8% of the retired did. In this study, as in Wells & Gubar, income was found to be related to FLC. Income rises until stage 5; then it drops. FLC was related to ethnic background in that 36% of Anglo-Americans were in stages 6, 7, and 8, whereas only 22% of Mexican-Americans were in these stages. The extended families of Mexican-Americans clearly influence this relationship.

FLC is a quite robust measure for capturing the variance of the demographic variables. As a general measure of influence it is statistically useful and analytically meaningful. Only one variable, area type, was not related to FLC. Finally, it appears that FLC is distributed similarly in rural, urban, and suburban areas.

Frequency of Recreation

Table 2 presents the hours per week spent in recreation away from home by FLC. As might be expected, bachelors spend the greatest amount of time in recreation away from home. Likewise, over half the newly married couples without the responsibility of children spend more than 5 hours per week recreating. The frequency patterns of stages 3 and 4 are quite similar in that there is a slight drop in the greater than five hour per week categories. The most dramatic change in recreation frequency occurs between stages 4 and 5. Fifty-four percent of the full nest II stage recreate less than 2 hours per week away from home. This might be attributable to the fact that the youngest child at home is older than 14. Thus, the children are becoming responsible for their own recreation lessening the

TABLE 1  
FLC DESCRIPTION & OPERATIONAL DEFINITIONS

Stage	Title	Description	Operational Definition	N
1.	Bachelor	young, single	Single, <40, no kids	43
2.	Newly Married	young, married, no kids	Married, <40, no kids	47
3.	Full Nest I (a)	young, married, youngest < 6	Kids; youngest < 6 <sup>a</sup>	151
4.	Full Nest I (b)	young, married, youngest ≥ 6	Kids; youngest ≥ 6, <14	142
5.	Full Nest II	older, married, dependent kids	Kids; youngest ≥ 14, at home	59
6.	Empty Nest (a)	older, married, no kids at home, head working	Married, ≥ 40, no kids at home; works	71
7.	Empty Nest (b)	older, married, no kids at home, head retired	married, ≥ 40, no kids at home, retired	79
8.	Solitary Survivors (a)	older, single, working	Single, ≥ 40, works <sup>b</sup>	55
9.	Solitary Survivors (b)	older, single, retired	Single, ≥ 40, retired	
				647 <sup>c</sup>

a. Variable measured if children were preschool, grade school, or high school.

b. Stages 8 and 9 were combined due to small cell size.

c. Three respondents failed to answer one of the classifying questions.



frequency of family type outings. This recreation frequency pattern appears to continue through stages 6, 7, and 8. In these stages, about 1/3 of respondents

stated they did not recreate away from home. Overall, Table 2 shows definite patterns of recreation shifts over the family life cycle.

TABLE 2  
PERCENT OF FLC STAGE SPENDING HOURS PER WEEK RECREATING AWAY FROM HOME

FLC Stage	Hours Per Week					Total
	0	Up to 2 hrs	2 to 5 hrs	5 to 10 hrs	10 & over	
1. Bachelor	7	9	9	35	40	100%
2. Newly Married	15	17	13	23	32	100%
3. Full Nest I (a)	10	20	28	26	16	100%
4. Full Nest I (b)	14	15	22	33	16	100%
5. Full Nest II	25	29	24	14	8	100%
6. Empty Nest (a)	34	18	17	17	14	100%
7. Empty Nest (b)	33	15	13	20	19	100%
8. Solitary Survivors	33	16	17	18	16	100%

#### Kinds of Leisure

Table 3 shows the percentage participating in 18 different leisure behaviors for each stage in the family life cycle. All activities in Table 3 were significant at the .05 level. Two sports (golf and shuffle board) are not included because participation rate did not differ by FLC. For each stage, a profile of activity usage can be developed. For example, the bachelor stage was a heavier than average user of tennis, lake swimming, volleyball, horse riding, hiking, camping, running, and dancing. The bachelors were lighter than average in hunting and fishing. The newly married show less activity than stage 1, but have recreation profiles similar (except for hunting and fishing) to bachelors. The family life cycle stage 3 shows about average usage on all activities. This slow down of recreation behavior is in keeping with the general notion that the family concentrates on nest building activities in stage 3. This is shown for most of the activities with the possible exceptions of dancing and lake swimming. As one might expect, the respondents in stage 4 continued to report strong usage of family type activities (i.e., picnic and pool swimming) and those behaviors that can be done by subsets of the family (biking by children). Although FLC stage 5 showed a strong tendency to recreate less than stages 1 to 4, certain behaviors remain strong -- tennis, soccer (children), water skiing horseback riding, hiking, biking, and hunting. Although there is a tendency for most activities to be reduced through stages 6, 7, and 8, some behaviors remain relatively strong across the three groups -- swimming, picnic, hiking, camping, fishing, and dancing. The combined effect of examining Tables 2 and 3 shows that FLC does reveal patterns in both the frequency and types of recreation behavior favored in each stage of the life cycle.

#### Public Park Usage

The second research question addressed in this study referred to the usage pattern for public facilities. Table 4 shows the number of times per year the household used public parks. By far, the heaviest users of

parks are bachelors and families with children between ages 6 and 14. Sixty-eight percent of the bachelors and 67% of full nest I (b) used a park 10 times or more per year while only 53% and 60% of the families in stages 2 and 3 reported that level of park usage. As can be seen from Table 4, the percentages for 10 or more park uses per year drops off dramatically after stage 5. Given that parks are designed to provide recreation with little need for travelling long distances, one might argue that since over 70% of stages 7 and 8 did not use a park more than 3 times a year, that the reduced general level of leisure activity is manifested by little park usage. Another explanation could be that the Parks Department of Pima County is not offering the activities that attract people from stages 7 and 8.

#### Adequacy of Facilities

When "don't know" responses were included in the analysis, all evaluative questions produced significant results. Since "don't knows" about various age group programs are related to FLC, they were deleted from the analysis. Thus, the evaluative results presented here are from respondents familiar with the activity.

Table 5 presents the results of a neighborhood availability question. It should be noted that those stages without children tend to rate the availability in neighborhoods as fair or poor with less frequency than the stages that have children to consider. Thus, perceptions of facility availability is related to the stage in the FLC. However, a similar question to the one shown in Table 5 asking about how "adequate" the facilities are in the area produced insignificant results.

When the questions were more specific, significant differences were noted for different programs. Table 6 summarizes the results of a program by program rating. The respondents were asked if the program was excellent, good, fair, or poor. Overall significance in these type analyses are secondary to examining percentages for different FLC groups. For example, the lack of

TABLE 3  
PERCENT PARTICIPATING IN CERTAIN RECREATION ACTIVITIES BY FLC

FLC Stage	Recreation Activities																	
	Tennis	Hand-ball	Swim lake	Swim pool	Base-ball	Soc- cer	Vol- ley ball	Skeet trap	Pic- nic	Horse ride	Hike	Water ski	Bike	Camp	Hunt	Fish	Run	Dance
1. Bachelor	58	23	37	74	28	7	44	5	79	26	74	9	56	70	12	19	42	26
2. Newly Married	43	26	19	49	40	6	32	19	79	17	47	9	43	60	23	40	17	6
3. Full Nest I (a)	38	16	23	52	29	7	23	8	79	18	33	5	39	44	21	29	13	11
4. Full Nest I (b)	46	16	24	76	34	10	24	16	83	25	46	11	59	56	30	45	11	13
5. Full Nest II	58	19	17	63	36	10	15	12	63	22	42	12	54	49	29	36	5	9
6. Empty Nest (a)	23	9	1	24	13	1	10	9	42	11	23	4	17	24	16	20	6	10
7. Empty Nest (b)	5	0	3	24	0	0	3	4	49	4	11	0	18	22	10	23	1	6
8. Solitary Survivor	9	4	11	22	7	4	0	6	58	7	18	0	9	15	9	15	4	11
Overall	35	13	18	50	24	6	19	10	69	17	35	7	39	42	20	30	11	11
Range	53	26	36	54	40	10	44	15	41	22	63	12	50	55	21	30	41	20

TABLE 4  
NUMBER OF TIMES/YEAR HOUSEHOLD USED PARKS

FLC Stage	Times Per Year						Total
	0	1-3	4-9	10-20	21-61	Over 61	
1. Bachelor	9	19	14	19	21	28	100%
2. Newly Married	15	13	19	15	19	19	100%
3. Full Nest I (a)	7	13	20	19	24	17	100%
4. Full Nest I (b)	5	13	15	18	19	30	100%
5. Full Nest II	19	13	17	19	8	17	100%
6. Empty Nest (a)	34	20	15	15	10	6	100%
7. Empty Nest (b)	52	20	9	9	6	4	100%
8. Solitary Survivor	38	33	9	4	7	9	100%
Overall	19	17	16	15	16	17	

significance for preteen programs is not as critical as the fact that 54% of the sample rated the program as "poor" and more than half of the respondents were from the most affected groups -- stages 3 and 4. Similarly, the significant Chi-Square for senior citizens programs is of secondary importance to the fact that nearly half of those in stages 6 to 8 rated these programs high, but 70% of the lower ratings came from those not directly using such activities.

Thus, the FLC concept helps to provide a measure of interpretability to program evaluations in that FLC

stages serve as easily identifiable target groups in the community to whom certain recreation activities are supposed to be useful. The answer to the third and fourth research questions appears to be affirmative. The FLC concept offers a viable way to analyze perceptions and recreation behavior within a community.

#### CONCLUSION

Although some applications of life cycle have not been as fruitful as expected (Jain 1975), the present study

showed that FLC is a promising independent variable for future leisure research. Municipal officials concerned not with personal income, but with the public's needs and usage patterns so as to evaluate and design different programs, should use FLC in demand and effectiveness analyses. FLC offers a construct that could provide a

rich independent variable to analyze present and anticipated recreation needs within the community. Likewise, FLC offers a means of segmenting markets into target groups whose recreation needs and subsequent behaviors are relatively homogeneous.

TABLE 5  
RATINGS ON AVAILABILITY IN NEIGHBORHOOD BY FLC<sup>a</sup>

FLC Stage	Availability Rating				Total
	Excellent	Good	Fair	Poor	
1. Bachelor	17%	27%	25%	31%	100%
2. Newly Married	7%	26%	12%	55%	100%
3. Full Nest I (a)	7%	26%	15%	52%	100%
4. Full Nest I (b)	16%	21%	12%	51%	100%
5. Full Nest II	7%	22%	13%	58%	100%
6. Empty Nest (a)	9%	37%	6%	48%	100%
7. Empty Nest (b)	17%	40%	6%	37%	100%
8. Solitary Survivor	16%	32%	16%	36%	100%
Overall	12%	28%	13%	48%	

<sup>a</sup> N=609 Don't knows were excluded.

TABLE 6  
RESULTS OF RATING OF DIFFERENT PROGRAMS BY FLC

<u>Program</u>	<u>FLC Significance</u>	<u>Comments</u>
Preteen	No Significance	Programs were rated poor by 54% of the sample with little FLC stage variation. But 57% of the poor ratings were from stages 3 and 4.
Teenage	No Significance	Poor rating given by 58% of sample.
Adult	No Significance	Good rating by 35% of sample, but 49% of Empty Nest (a) rated it as good.
Senior Citizens	$X^2, p > .01$	48% of excellent or good ratings came from stages 6 to 8. 70% of fair or poor ratings came from stages 1 to 4.
Outdoor Facilities	$X^2, p > .05$	53% of fair or poor ratings came from stages 3 and 4.
Indoor Facilities	$X^2, p > .001$	Strong poor ratings from stages 3 and 4. Excellent ratings from stages 7 and 8.
Cultural Activities	$X^2, p > .001$	Stages 3, 4, and 5 rated these programs poor.

#### REFERENCES

James F. Engel, David T. Kollat, and Roger D. Blackwell, Consumer Behavior, (Hinsdale, Illinois: Dryden Press, 1978)

Subhash C. Jain, "Life Cycle Revisited: Applications in Consumer Research," In M. Schlinger (Ed.) Advances in Consumer Research, Volume 2, 1975, 39-49.

Glenn S. Omura and W. Wayne Talarzyk, "Relationships Between Consumers' Shopping and Leisure Activities and Their Attitudes Toward the Energy Crises: A Cross Sectional Study," In M. Schleinger (Ed.) Advances in Consumer Research, Volume 2, 1975, 803-815.

Fred D. Reynolds and William D. Wells, Consumer Behavior, (New York: McGraw-Hill, 1977)

Justin L. Voss and Roger D. Blackwell, "Markets for Leisure Time," In M. Schlinger (Ed.) Advances in Consumer Research, Volume 2, 1975, 837-845.

William D. Wells and George Gubar, "Life Cycle Concept in Marketing Research," Journal of Marketing Research, (1966) 355-63.

## SOCIAL CLASS DETERMINANTS OF LEISURE ACTIVITY

Robert B. Settle  
Pamela L. Alreck  
Michael A. Belch  
San Diego State University

### Abstract

A survey of 975 metropolitan West Coast adults from the general public obtained participation rates for 100 leisure activities grouped as entertainment, competitive sports, non-competitive sports, and hobbies. Participation was cross-tabulated with five socioeconomic variables: self-rated social class, education, income, occupational class and occupational group, and with five demographic variables: age, sex, marital status, employment status, and family life cycle. The results, reported as the trend in participation rate across categories, revealed that (a) the demographics were better predictors than the socioeconomic factors, (b) both groups are systematically related to leisure behavior, and (c) education was by far the best single socioeconomic determinant of leisure choices.

### Introduction

Social class status has long been thought to influence the members' behavior. Casual observation of different social classes leads to the intuitive belief that behavior patterns differ from one class to the next, and several studies have actually detected such differences. Yet there are many confusing issues surrounding both the theoretical construct of social class and the behavior patterns that are supposed to be affected by social status. There is general agreement that social class membership is determined by a person's status on several socioeconomic variables but no clear consensus about what variables must be included or what weight should be placed on each. There are also problems in identifying the behavioral effects of class membership. Even though preferences may differ by social class, the individual's actions may be constrained by situational factors that are not related to social status. The objective of this project was to measure the effect of four variables related to social class on people's leisure behavior.

### Social Class and Consumer Behavior

Attempts to relate social class differences to specific consumer behaviors have resulted in less than conclusive results. Martineau (1958) presented information supporting his contention that social class affects perceived risk, choice making, and store selection, among other factors. Levy (1966) has demonstrated differences between classes in respect to values, interpersonal attitudes, shopping behaviors, and media selection, among others. Mathews and Slocum (1969) established social class differences in respect to credit card usage. Rich and Jain (1968), on the other hand, noted little, if any differences between classes in respect to fashion interests, sources of shopper information, interpersonal influences, or other factors influencing shopping behavior.

A number of hypotheses can be formulated to account for the lack of consistent findings of these studies, as well as many others attempting to differentiate behavior patterns on the basis of social class. Inconsistent definitions of social class, confounding influences such as status inconsistency, and the changing structure of American society are but a few factors that may intervene in these studies. This study incorporates a mea-

sure of self-rated social class membership with educational status, family income, and two methods of categorizing occupational status, together with the demographic variables of age, sex, marital status, and family life cycle stage. This permits a comparison of the relative potency of the different variables in shaping behavior. The recency of the study may also account for the changing social class structure.

### Social Class and Leisure Activities

Most definitions of leisure establish it as time free from work or work related demands. Leisure time is, then, discretionary time, and leisure activities are a matter of individual judgment and undirected choices (Voss, 1967, p. 101). As such, leisure should provide a much better indication of differences in preferences that may exist among social strata. Substantiation for this position can be found in the sociological literature where a number of studies have examined the effects of social class and occupational status on the selection of leisure activities. Without attempting to review all of the literature in this area, a sampling of the studies will provide a perspective.

Reisman (1954), White (1955), Clark (1956), Gerstly (1961), and Burdge (1969) report similar findings, essentially demonstrating that differences in social class and/or occupational milieu are responsible for differences in preferred leisure activities. The conclusions of these studies suggest that occupational positions lead to differences in leisure activities within strata, that as prestige increases, involvement in the variety of social activities also increases in a linear relationship, and that individuals are more likely to engage in activities perceived as consistent with their social standing. MacDonald *et. al.* (1949) provided supporting evidence contending that children use leisure to prepare for further upward mobility.

Studies by Kaplan (1960), Cunningham *et. al.* (1970), Murphy (1974), and Kelly (1975), using the same basic paradigm, provide contrary information to the studies just reported. These studies found that neither social class nor occupational prestige was an effective indicator of leisure preference. In summary, the conflicting results of these studies prohibits the establishing of firm conclusions. One possible explanation is evident, and that is the dynamic nature of contemporary American social structure. The majority of early studies report differences among social classes, while more recent studies provide findings of homogeneity among strata. This study may provide additional insight into the existence or absence of class differences in behavior today, based on a very recent measure of current behavior patterns.

### Status Inconsistency

Classification of an individual into a given social strata is relatively easy and unambiguous when several of the determining variables are all consistent with one another. For example, a person whose income, wealth, education, and prestige of occupation were all at about the fifth percentile for that society could be classified as lower class without much equivocation. It also

seems likely that the individual's self-rating of social class would be to that category. When, on the other hand, the variables regarded as determining social class membership differ widely from one another, this status inconsistency may make classification difficult, if not impossible. How should a person whose income is in the sixtieth percentile and whose occupational prestige is in the tenth percentile be classified? Even if weights are established and a weighted average used for classification, the variance in status from one variable to the next may tend to blur the person's own perceptions of class membership. Such an individual might not be expected to behave in the same manner as another member of the same strata whose strata on the determining variables are identical.

Such factors as the increase in American affluence, the progress of labor organization resulting in higher wages for the working class, and other similar changes in the socioeconomic system may have increased status inconsistency to the point where the conglomerate concept of "social class" has less meaning than more unitary social status variables. Examination of the various forces leading to status inconsistency is beyond the scope of this project, however, the relative influence of some variables commonly regarded as determining social class can be measured empirically.

#### Hypotheses

- I. Patterns of leisure behavior will be systematically related to self-rated social class, income, education, and occupational status.
- II. Education will exert the strongest influence on leisure activity and income will exert the weakest influence, of the individual variables.

#### Method

To test the hypotheses, a survey questionnaire was individually self-administered to a convenience sample of about one thousand metropolitan West Coast adults from the general public. Field workers were assigned data collection quotas based on age, sex, and occupation to insure that the demographic distributions of the sample would correspond to census data for the same geographic region.

#### Questionnaire

Each data collection instrument consisted of a cover letter explaining the nature of the project, a sixty-four item, trait-specific personality test, a series of one hundred items about leisure activities, ten common pastimes, twenty-nine questions about habits of time use, and eleven demographic items. The personality test and items relating to time patterns are the object of another study and are not reported here.

The one hundred leisure activity items were divided into four groups of twenty-five each, and classified as Entertainment, Competitive Sports, Non-Competitive Sports, and Hobbies, respectively. The activities that were the content of the items are shown in Tables 1 through 4, in the following section. Respondents were asked how many times per week, month, or year they participated in each of the activities in the first three categories. For the hobbies, they indicated the number of hours per day, week, or month they engaged in each, and this same measure was obtained for the ten pastimes listed. The demographic items were listed by category, except for education, occupation and income. Respondents reported the last year of school they completed and these data were later classified into categories. They filled in their occupational title and a phrase describing what they did on the job, and these data were later classified into

occupational class and group categories. Approximate annual family income was recorded on an optional basis in thousands of dollars, and these too, were converted to categories for analysis. Lastly, each respondent was classified into a stage of the family life cycle, based on marital status, age of youngest child, age of respondent, and employment status.

#### Administration

Nine hundred ninety-six questionnaires were equally divided among eighty-three student field workers who collected the data as part of a Consumer Behavior class project. Each received about two hours of instruction on the data collection and instrument prior to going into the field, and each had completed the brief psychological test included in the questionnaire and received their own individual, six-page, computer-generated interpretations. The field workers only tasks were to select respondents, enlist their cooperation, deliver and retrieve the questionnaire, and transfer the data to code sheets for optical scanning. Instructions within the questionnaire were addressed to the respondent and the instrument was individually self-administered. As an inducement to participate, the field workers explained the personality test to the potential respondent and each was provided with his or her own interpretations within a few weeks after completing the questionnaire. These interpretations, mailed directly to the respondents' homes, provided one hundred percent validation of the data collection.

Of the 996 questionnaires sent into the field, 984 were recovered. Seven questionnaires were rejected because they were not substantially complete, and two others were lost due to errors in data transfer, leaving a total sample size of 975 respondents.

#### Analysis

Statistical analysis for this report focused only on the 100 leisure activities and the social class and demographic variables. The data description for the independent variables is provided in the form of frequency and percentage distributions of response, to reveal the nature of the sample obtained. Leisure activity items were converted to dichotomous variables according to whether or not the respondent had participated in the activity. These were then cross-tabulated with the discrete social class and demographic items and the Chi square statistic used to measure statistical significance at the .05 alpha level.

While sacrificing much of the detail contained in the data, the procedure for analysis serves to condense and abbreviate the results for clarity of presentation. Even so, there remained exactly 1,000 relationships to portray. Time and space limitations demanded a novel yet easily understood format. Each of the contingency tables was checked for statistical significance. For those tables which were significant, the percentage of each category that reported participating in the activity was noted.

In this report, the tables depicting these relationships are in matrix form with the rows defined by the various leisure activities within one category and the columns defined by either a social class or a demographic variable. In the body of the table, each cell represents the relationship between one dependent and one independent variable. The content of the cell indicates (a) whether or not the relationship was statistically significant at the .05 alpha level, and if so, (b) the direction of the relationship. If the relationship between the social class variable and the participation in the activity is not significant, the cell is blank. If the relationship is significant, the cell contains a single

line.

The direction of the relationship between the dependent and the independent variable for a given cell is shown by the slope of the line. If the rate of participation in the activity tended to increase for higher levels of the social class or demographic variables, the line has a positive slope. If the rate of participation tended to decrease as the independent variable ascended, the line has a negative slope. For those contingency tables where the rate of participation fluctuated significantly across the categories of the independent variable but no clear and consistent trend was discernible, the cell contains a horizontal line. The six categories of employment status defy arrangement into a hierarchy, and consequently, a horizontal line was used to indicate a significant relationship without regard for direction of trend. The other independent variables ascend in the order in which the categories are presented for data description, from lowest to highest income, education, age, or family life cycle, etc. Sex ascends from male to female and marital status from single to married (though the arrangement is admittedly an arbitrary choice). The use of this unique report format permits portrayal of both the number of significant relationships and the direction of the relationships. Comparison of columns reveals the relative potency of each independent variable, and inspection of rows shows the nature of activities that are influenced.

### Results

The direction of the significant relationships between social strata and demographic variables, on the one hand, and the dependent variables of participation in the various leisure activities, on the other, are shown in Tables 1 through 4. The activities classified as entertainment, competitive sports, non-competitive sports, and hobbies will be discussed in turn.

#### Entertainment Participation

Inspection of Table 1 indicates that the respondents' level of education was significantly related to participation in 20 of the 25 entertainments listed. The rate of participation increased with education for 18 of the 20 significant items, while visits to a nightclub, lounge or bar appeared to increase to middle levels of education, then decrease for higher levels. The relationship between education and visiting a swap meet was inverse; the higher the level of education, the less likely the respondent would participate.

Eleven of the 25 entertainment items provided a significant relationship to self-rated social class and all but one of the 11 were also systematically related to education. The direction of the relationships were identical for all but one; going out to dance was more often reported for lower and higher class respondents and less often reported for middle class people. The two occupational categories were identical in their patterns of systematic relationships with leisure entertainment activities, except that visits to swap meets were related to occupational groups but not to occupational classes. This seems to indicate that, for entertainment activities, the two-category classification of white collar/blue collar proved to be as indicative as the finer increments: semi- and unskilled, skilled labor, technical and clerical, or professional. A total of 10 of the activities were significantly related to occupational groups and 9 to occupational class.

Income was systematically related to participation in only 8 of the entertainment activities listed. The only leisure activity relationship uniquely related to income was visits to a card room, casino or bingo game. There

were 57 significant relationships between social strata variables and participation in leisure activities classified as entertainment (while only about 6 or 7 would be expected by chance at the .05 alpha level). These results furnish support for the first hypotheses, that social stratification would be systematically related to participation in leisure activities. The second hypothesis, that education will provide the best indication of leisure activity and income the least, was also supported by the results.

TABLE 1

#### RELATIONSHIPS BETWEEN SOCIAL STATUS AND ENTERTAINMENT

Leisure Activity	Soc. Cls	Income	Education	Occ. Cls.	Occ. Group	Age	Sex	Mar. Status	Emp. Status	Life Cycle
Go to the movies . . . . .	/	/	/	/	/	/	/	/	/	/
Go to a concert . . . . .	/	/	/	/	/	/	/	/	/	/
Visit a zoo or public park	/	/	/	/	/	/	/	/	/	/
Go to the library . . . . .	/	/	/	/	/	/	/	/	/	/
Visit a museum or art gallery . . . . .	/	/	/	/	/	/	/	/	/	/
Go to the beach . . . . .	/	/	/	/	/	/	/	/	/	/
Go to dinner at a restaurant . . . . .	/	/	/	/	/	/	/	/	/	/
Go for a drive in the car .	/	/	/	/	/	/	/	/	/	/
Go to a play . . . . .	/	/	/	/	/	/	/	/	/	/
Attend a lecture . . . . .	/	/	/	/	/	/	/	/	/	/
Go to a ballet . . . . .	/	/	/	/	/	/	/	/	/	/
Go on a picnic . . . . .	/	/	/	/	/	/	/	/	/	/
Go to a party . . . . .	/	/	/	/	/	/	/	/	/	/
Go out to dance . . . . .	/	/	/	/	/	/	/	/	/	/
Visit a nightclub, lounge or bar . . . . .	/	/	/	/	/	/	/	/	/	/
Go to an opera . . . . .	/	/	/	/	/	/	/	/	/	/
Go for a walk . . . . .	/	/	/	/	/	/	/	/	/	/
Visit a club, lodge, social organization . .	/	/	/	/	/	/	/	/	/	/
Attend non-worship religious event . . . . .	/	/	/	/	/	/	/	/	/	/
Go window shopping as recreation . . . . .	/	/	/	/	/	/	/	/	/	/
Visit a swap meet . . . . .	/	/	/	/	/	/	/	/	/	/
Go to a spectator sports event . . . . .	/	/	/	/	/	/	/	/	/	/
Visit a card room, bingo, casino . . . . .	/	/	/	/	/	/	/	/	/	/
Go to a card party . . . . .	/	/	/	/	/	/	/	/	/	/
Go out visiting friends, relatives . . . . .	/	/	/	/	/	/	/	/	/	/

Of the demographic items, age was significantly related to 21 of the entertainment activities, as was family life cycle. The sex of the respondent was systematically related to only seven activities. In total, there were 72 significant relationships between demographic items and entertainment activity participation.

Competitive Sports Participation

The significant relationships between social strata and demographic variables and competitive sports participation are shown in Table 2. Only 21 of the 25 sports activities are shown because the last four items were for "other" activities to be written in by respondents. None provided a sufficient frequency for contingency table analysis.

TABLE 2

RELATIONSHIPS BETWEEN SOCIAL STATUS AND COMPETITIVE SPORTS

Leisure Activity	Soc. Cls.	Income	Education	Occ. Cls.	Occ. Group	Age	Sex	Mar. Status	Emp. Status	Life Cycle
Baseball or softball . . .	/	/	/	/	/	/	/	/	/	/
Basketball . . . . .	/	/	/	/	/	/	/	/	/	/
Football . . . . .	/	/	/	/	/	/	/	/	/	/
Soccer . . . . .	/	/	/	/	/	/	/	/	/	/
Track and field sports . .	/	/	/	/	/	/	/	/	/	/
Racquetball . . . . .	/	/	/	/	/	/	/	/	/	/
Paddleball or squash . . .	/	/	/	/	/	/	/	/	/	/
Handball . . . . .	/	/	/	/	/	/	/	/	/	/
Tennis . . . . .	/	/	/	/	/	/	/	/	/	/
Golf . . . . .	/	/	/	/	/	/	/	/	/	/
Bowling . . . . .	/	/	/	/	/	/	/	/	/	/
Volleyball . . . . .	/	/	/	/	/	/	/	/	/	/
Badminton . . . . .	/	/	/	/	/	/	/	/	/	/
Skeet or target shooting .	/	/	/	/	/	/	/	/	/	/
Archery . . . . .	/	/	/	/	/	/	/	/	/	/
Table tennis or ping pong . . . . .	/	/	/	/	/	/	/	/	/	/
Chess . . . . .	/	/	/	/	/	/	/	/	/	/
Backgammon . . . . .	/	/	/	/	/	/	/	/	/	/
Other "board" games . . . .	/	/	/	/	/	/	/	/	/	/
Card games . . . . .	/	/	/	/	/	/	/	/	/	/
Pool or billiards . . . . .	/	/	/	/	/	/	/	/	/	/

The pattern of relationships revealed by Table 2 indicates that the demographic items were much more strongly related to participation in these sports activities than were the social strata items, by a ratio of 87 to 37. Nearly half of the significant associations were with education, while both income and occupational group each accounted for only 5, and occupational class was the least effective predictor of activity with 4 systematic relationships to the 21 dependent variables. As with the previous table, these results provide support for both of the hypotheses.

Non-Competitive Sports Participation

Table 3 contains the report of the association between participation in non-competitive sports and respondent status on the demographic variables and those regarded as determinants of social status. As with competitive sports, age, family life cycle, and employment status were strongly related to participation. There were 37 significant relationships between social status varia-

bles and the activities. Of those, 18 were with education, 8 with income, 5 with occupational group and 3 each with self-rated social class and occupational class.

TABLE 3

RELATIONSHIPS BETWEEN SOCIAL STATUS AND NON-COMPETITIVE SPORTS

Leisure Activity	Soc. Cls.	Income	Education	Occ. Cls.	Occ. Group	Age	Sex	Mar. Status	Emp. Status	Life Cycle
Fishing and casting . .	/	/	/	/	/	/	/	/	/	/
Sailing . . . . .	/	/	/	/	/	/	/	/	/	/
Power boating and water skiing . . . . .	/	/	/	/	/	/	/	/	/	/
Swimming and diving . .	/	/	/	/	/	/	/	/	/	/
SCUBA diving and snorkeling . . . . .	/	/	/	/	/	/	/	/	/	/
Surfing . . . . .	/	/	/	/	/	/	/	/	/	/
Walking, jogging and running (city). . . . .	/	/	/	/	/	/	/	/	/	/
Hike, backpack, climb (wilderness). . . . .	/	/	/	/	/	/	/	/	/	/
Camping and hunting . .	/	/	/	/	/	/	/	/	/	/
Flying a private plane .	/	/	/	/	/	/	/	/	/	/
Gliding and parachuting	/	/	/	/	/	/	/	/	/	/
Skateboarding and roller skating . . . . .	/	/	/	/	/	/	/	/	/	/
Ice skating . . . . .	/	/	/	/	/	/	/	/	/	/
Snow skiing . . . . .	/	/	/	/	/	/	/	/	/	/
Sledding, toboggan, snowmobiling . . . . .	/	/	/	/	/	/	/	/	/	/
Weight lifting and calisthenics . . . . .	/	/	/	/	/	/	/	/	/	/
Martial arts (Karate, Judo, etc.) . . . . .	/	/	/	/	/	/	/	/	/	/
Playing "catch" or frisbee . . . . .	/	/	/	/	/	/	/	/	/	/
Horseshoes, yard darts, croquet . . . . .	/	/	/	/	/	/	/	/	/	/
Bicycling . . . . .	/	/	/	/	/	/	/	/	/	/
Motorcycling (tour or off-road) . . . . .	/	/	/	/	/	/	/	/	/	/
Dune-buggy or off-road driving . . . . .	/	/	/	/	/	/	/	/	/	/
Riding and jumping horses . . . . .	/	/	/	/	/	/	/	/	/	/

For both competitive and non-competitive sports, grouping occupations into four categories rather than just two provided only very slightly more prediction of participation in the activities. The number of significant relationships shown on Table 3 provides support for Hypothesis I. The second hypothesis received partial support in that education was the best predictor of behavior, however income was not the least potent, as hypothesized. Examination of non-competitive sports items reveals that many require expensive equipment or fees, which may account for the relationships to income. This was not the case for competitive sports included in the previous table.



Participation in Hobbies

The systematic relationships between participation in 22 hobbies and the independent variables are shown in Table 4. As with the others, the demographic variables provided more influence on participation than did the items associated with social status. In all, there were 17 significant relationships between participation and the social status measures (while only 5 or 6 would be expected purely by chance at the .05 level of significance). Only one hobby, clothing design and sewing, was systematically related to self-rated social class. Collecting things was related to income, and adult education participation was significantly related to occupational class, while four hobbies were related to occupational group, and four hobbies were related to occupational class, while four hobbies were related to occupational group.

TABLE 4

RELATIONSHIPS BETWEEN SOCIAL STATUS AND HOBBIES

Leisure Activity	Soc. Cls.	Income	Education	Occ. Cls.	Occ. Group	Age	Sex	Mar. Status	Emp. Status	Life Cycle
Cooking and food preparation . . . . .		^								
Clothing design and sewing . . . . .	^									
Painting or drawing . . . . .		^								
Photography . . . . .		^								
Interior decorating . . . . .		^								
Model building . . . . .										
Crafts, ceramics or macrame . . . . .		^								
Knit, crochet, weave, needlepoint . . . . .										
Woodworking, furniture finishing . . . . .										
Coin or stamp collecting										
Collections of other items . . . . .		^								
Musical instrument, voice practice . . . . .		^								
Reading fiction or non-fiction . . . . .										
Adult education classes . . . . .										
Gardening, yard care, houseplants . . . . .										
Auto care, mechanical projects . . . . .										
Model building, racing, showing . . . . .										
Antique hunting, finish, collecting . . . . .										
Electric or electronic projects . . . . .										
Ham or CB radio and communications . . . . .										
Creative writing . . . . .										
Drama and performing arts . . . . .										

As with the other categories of leisure activity, education was by far the best single predictor, significantly associated with 10 of the hobbies. Income shared the weakest position with self-rated social class and occupational class. These results provide support for both of the hypotheses.

Summary of Results

Considering only those independent variables associated with social status, the educational level of the respondent was by far the most effective determinant of leisure activity, with significant relationships with 65 of the 91 activities listed. In descending order of effect, the other variables were occupational group with 24 relationships, income with 22, self-rated social class with 21, and occupational class with the least number, 17. The total number of significant associations was 149 of a possible 455. At the .05 alpha level, one would expect only about 23 to be significant purely by chance. These results provide support for the first hypothesis, that the patterns of leisure activity will be significantly influenced by variables associated with social status. It was also hypothesized that education would provide the best single predictor of participation in the leisure activities, while income would be the least effective predictor. This hypothesis received only partial support. While education was the best predictor, both occupational class membership and self-rated social class proved to have fewer relationships to the dependent variables than did income.

Examination of the direction of the relationships indicates that for the vast majority of activities, participation tends to be less for lower levels of social status and greater for higher levels. There were a minority of activities which showed a curvilinear relationship, where participation increased from lower to middle levels of social status, then decreased for the highest levels. There were a few activities, such as pool or attending a swap meet, that were more typical of lower levels than of higher levels of social status, but these were intuitively understandable.

Conclusions

The patterns of relationships between participation in leisure activities and the independent variables are very consistent with prior expectations in many respects. For example, demographic factors of age, sex, and family life cycle provide significant relationships with those leisure pursuits that demand strenuous physical activity as one would expect. These kinds of observations about the results are, of course, only "face" validity, but they do encourage the belief that the measurements are reliable and valid. Face validity is unimportant when present but very important when it is absent.

Education versus Income and Occupation

The major objective of this study was to measure the effect of four variables indicative of social status on consumers' selections of leisure activity. The study also sought to compare the relative potency of each in shaping behavior. Education proved to be the most influential variable, by far. Why might this be so?

In pursuit of the answer to this question, one might logically ask how education differs from the other variables. One very fundamental difference lies in the distinction between an indicator of social status and a determinant of social status. If a variable is an indicator, this implies only a correlation, but if it is a determinant, that implies causality. For example, in many communities the size of a person's front lawn is an excellent indicator of social status, but this is

only an indicator and not a determinant. If a person from a high social station were to acquire a home with a very small front lawn, he would not sacrifice social status, he would only decrease the predictability of the indicator.

Self-rated social class status can be declared an indicator and not a determinant of social class without much danger of conflict. Few would argue that a self-perception of a particular social class membership would, in fact, propel the person to that particular station. This is not the case, however, for income and occupation or occupational prestige. These are often regarded as true determinants of social status. The contributions upon which this society is supposed to accord social status often result from one's occupation, and the relative value of the contribution is thought to be denominated in terms of monetary rewards, or income.

But what happens when income deviates from an accurate reflection of social contribution? Both income and the occupation providing it might remain as indicators of social class, out of inertia, however they would no longer be determinants of social status.

Several reasons can be cited for deviation of income from an accurate reflection of social contribution in our contemporary society. First, the socioeconomic system often uses income to compensate for occupations that are socially unattractive, such as when trash collectors are paid more than school teachers. Second, occupational organization and political power may force income disproportionate to the long-term social contributions of the occupation. Third, some vocations may provide very valuable social contributions but they may be so attractive in terms of intrinsic rewards that many will pursue them without demanding monetary rewards commensurate with contributions. To the degree that these conditions describe society today, income and occupation can be expected to diminish in importance as determinants of social status.

#### Education as a Direct Influence

To this point, it has been assumed that (a) social status influences consumer choices and (b) education is an accurate indicator or determinant of social class. Thus, the paradigm is one of a mediationist and the social class construct is an intervening variable between the educational process and overt behavior. An alternate perspective is that education directly affects choice behavior and at the same time, determines social status. (An even more radical view would be that education affects choices which, in turn, determine social status; an idea that Thorstein Veblen would have savored.) Regardless of one's choice of models, there is ample evidence that education is more of a developmental process than the adult pursuit of an occupation or the acquisition of income. We might expect the educational process to shape the tastes and values of the individual, directly changing preferences for the various forms of leisure activity. This proposition is not intended to deny the existence or effects of social stratification, but it does serve as a reminder that all influences on choice behavior are not mediated by social class membership.

#### Reconciliation and a Projection

The results of the earlier studies of social class and choice behavior indicated a significant relationship. Later studies were less successful in detecting a strong relationship. The results of this study indicate that there are probably still many differences among social strata; however, the measurement of the strata themselves should probably be based on variables other than income and occupational class or group, as in the past. It may prove fruitful in future research to first

explore the relative effects of several candidate variables in determining social status, selecting those that perform the best. Having done so, the second step in the process would be to define strata in terms of these factors. Lastly, the investigation of the effects class membership might be studied in terms of choice behavior. Research experience in the area of social stratification over the past several years, including this study, seem to offer this lesson: in a dynamic setting with an ever-increasing velocity of change, the clarity of yesterday's identities and relationships may fade very quickly into ambiguity and obsolescence.

TABLE A  
INDEPENDENT VARIABLE DATA DESCRIPTIONS

Category	Frequency	Percent	Adj. Percent
Age			
Teens	23	2.4	2.4
Twenties	215	22.1	22.1
Thirties	113	11.6	11.6
Forties	175	17.9	17.9
Fifties	242	24.8	24.8
Sixties	120	12.3	12.3
Seventies	87	8.9	8.9
Missing	0	0.0	---
Sex			
Male	485	49.7	50.0
Female	485	49.7	50.0
Missing	5	0.5	---
Marital Status			
Single	341	35.0	35.4
Married	621	63.7	64.6
Missing	13	1.3	---
Employment Status			
Company employed	427	43.8	45.1
Government employed	118	12.1	12.5
Self employed	131	13.4	13.8
Housewife	93	9.5	9.8
Unemployed	47	4.8	5.0
Retired	131	13.4	13.8
Missing	28	2.9	---
Family Life Cycle			
Bachelor	166	17.0	18.4
Newly Married	44	4.5	4.9
Full Nest I	52	5.3	5.8
Full Nest II	104	10.7	11.5
Full Nest III	145	14.9	16.1
Empty Nest - Employed	207	21.2	22.9
Empty Nest - Retired	69	7.1	7.6
Solitary - Employed	71	7.3	7.9
Solitary - Retired	44	4.5	4.9
Missing	73	7.5	---
Self-Rated Social Class			
Lower & Working Class	253	25.9	26.5
Middle Class	639	65.5	67.0
Upper Class	62	6.4	6.5
Missing	21	2.2	---
Income			
\$5,000 & Under	44	4.5	8.6
\$6,000 to \$10,000	76	7.8	14.9
\$11,000 to \$15,000	111	11.4	21.8
\$16,000 to \$20,000	85	8.7	16.7
\$21,000 to \$25,000	51	5.2	10.0
\$26,000 to \$30,000	44	4.5	8.6
\$31,000 to \$35,000	31	3.2	6.1
\$36,000 to \$40,000	18	1.8	3.5
\$41,000 & Over	49	5.0	9.6
Missing	466	47.8	---

Category	Frequency	Percent	Adj.Percent
<u>Education</u>			
Some Elem. School	23	2.4	2.4
Elem. School Only	24	2.5	2.5
Some High School	88	9.0	9.1
High School Grad.	294	30.2	30.5
Some College	279	28.6	29.0
College Graduate	140	14.4	14.5
Post Graduate	115	11.8	11.9
Missing	12	1.2	—
<u>Occupational Class</u>			
Blue Collar	363	37.2	43.5
White Collar	471	48.3	56.5
Missing	141	14.5	—
<u>Occupational Group</u>			
Semi- or Unskilled	178	18.3	21.3
Skilled Labor	185	19.0	22.2
Technical/Clerical	270	27.7	32.4
Professional	201	20.6	24.1
Missing	141	14.5	—

J. Voss, "The Definition of Leisure," Journal of Economic Issues, 1 (June 1967), 91-106.

R. C. White, "Social Class Differences in the Use of Leisure," The American Journal of Sociology, 61 (1955), 145-150.

#### References

- R. J. Burdge, "Levels of Occupational prestige and Leisure Activity," Journal of Leisure Research, 1 (Summer 1969), 262-274.
- A. C. Clarke, "The Use of Leisure and its Relation to Levels of Occupational Prestige," American Sociological Review, 21 (June 1956), 301-307.
- D. A. Cunningham, H. J. Montoye, H. L. Metzner, and J. B. Keller, "Active Leisure Activities as Related to Occupation," Journal of Leisure Research, 4 (Spring 1970), 104-111.
- J. E. Gerstle, "Leisure, Taste, and Occupational Milieu," in Erwin O. Smigel, ed., Work and Leisure. New Haven: College and University Press, 1963.
- M. Kaplan, "The Use of Leisure," in C. Tibbits, ed., Handbook of Social Gerontology. Chicago: University of Chicago Press, 1960.
- J. R. Kelly, "Life Style and Leisure Choices," The Family Co-ordinator (April 1975), 185-190.
- S. J. Levy, "Social Class and Consumer Behavior," in Joseph W. Newman, ed., On Knowing the Consumer. New York: John Wiley and Sons, 1966.
- M. MacDonald, C. McGuire and R. J. Havighurst, "Leisure Activities and the Socioeconomic Status of Children," American Journal of Sociology, 59 (May 1949), 505-519.
- P. Martineao, "Social Classes and Spending Behavior," Journal of Marketing, 23 (October 1958), 121-130.
- H. L. Mathews and J. W. Slocum, Jr., "Social Class and Commercial Bank Credit Card Usage," Journal of Marketing 33 (January 1969), 71-78.
- J. Murphy, "Leisure Determinants of Life Style," Leisure Today (November-December 1974), 3-5.
- L. Reismann, "Class, Leisure and Social Participation," American Sociological Review, 19 (February 1954), 76-84.
- S. V. Rich and S. C. Jain, "Social Class and Life Cycle as Predictors of Shopping Behavior," Journal of Marketing Research, 5 (February 1968), 41-49.

COMMENTS ON PAPERS ON LIFE CYCLE ANALYSIS

Robert Ferber, University of Illinois

As is evident from the papers, all of them deal with the life cycle concept in one way or another. Two of the papers attempt to emphasize the value of a life cycle concept for explaining consumer behavior, while the third includes a life-cycle variable among others in an attempt to assess the relative importance of these different variables in explaining one type of consumer behavior--recreation.

My approach in these comments will be to first make some remarks on each of the papers individually. Based on these evaluations, some comments will then be offered on the value of a life-cycle type of variable, and particularly on what needs to be done if the value of a life-cycle variable for analysis of consumer behavior is to be established. Also, given my awareness of the problems that exist with survey data of the type reported here, I have been unable to resist making a few final comments on how such data might be reported and interpreted when used for analytical purposes.

The Paper by Arndt

This paper tends to reemphasize the importance of the life cycle concept in influencing three types of activities--receipt of income, composition of consumption expenditures and the proportion spent on gifts (which are treated separately from other consumption expenditures). This is done by recourse to data from a national survey of consumer expenditures conducted in Norway in 1973, presenting breakdowns of these expenditures by stages of the family life cycle.

In line with these objectives, the paper presents a cross tabulation of the distribution of the 3,360 families interviewed in this survey by major expenditure categories by six stages in the life cycle, plus a seventh stage designated "other households". As in previous studies in this country and elsewhere, the results show that income (represented by total consumption expenditures) first rises over the life cycle and then falls; that the allocation of the family budget to different expenditures fluctuates in some predictable (but also some unpredictable) fashions; and that gifts as a percent of total expenditures tend to follow a U-shaped pattern.

Accordingly, do these results establish the value of the family life cycle concept in explaining spending patterns? I am afraid not. There are two types of problems with this analysis, one type associated with the analysis of the data and the other associated with the general approach.

The analysis of these data raises a number of questions, as follows:

1. The six stages of the life cycle used in this analysis account for only two-thirds of the families for which data were available. It seems to me that when over one-third of the observations in a national study have to be discarded because they do not fit a particular classification, something must be wrong with the classification. While the author points out one reason for this large loss, namely, that the upper age limit for dependent children was set at 16 by the Norwegian Central Bureau of Statistics, the fact remains that no attempt was made to deal with this problem. As the author points out in his concluding paragraph, living arrangements are changing, particularly in view of the increasing number of

divorces and remarriages. Hence, a life-cycle classification to be meaningful under present circumstances has to take these changes into account if it is to be useful for explaining consumer behavior.

2. The tendency to equate income with household expenditures is questionable, unless one wants to consider such expenditures as a measure of permanent income, something that is not discussed. Total expenditures and family income are undoubtedly highly correlated, as the author says, but there is no question that appreciable differences would result in expenditures, since saving is hardly likely to be constant over the life cycle.
3. From a survey point of view, the reported response rate of 71% in this survey raises questions about sampling biases, and also about the reliability of the data for life cycle analysis, neither of which is discussed. Especially noticeable, for example, is that the base for each of the two first stages of the life cycle is only about 100 families, or about 3% of the total sample. Both absolutely and percentage-wise, these are very small bases to use for analysis, especially in a nationwide study.
4. Also not receiving any consideration is the very likely possibility of response errors in the expenditure data. Expenditure studies of the type reported here are notorious for having such errors, with substantial proportions of the total known expenditures for a particular category being unaccounted for. Thus, an evaluation of the 1972-73 survey of consumer expenditures in this country by Robert Pearl of our staff, done for the U.S. Bureau of the Census, indicates that 30-40% of expenditures for meals away from home were not reported in the survey. These errors are by no means equally distributed by different types of families, and their consideration might alter substantially the nature of the inferences drawn from these data.
5. On a more detailed level, I am not convinced that the category "other goods and services" represents only restaurant meals and entertainment; such an "other" category very likely includes also a wide variety of other services (such as financial) and other types of expenditures that could hardly be classified as meals or entertainment. Their relative importance in this category is not mentioned, nor the very real possibility that they will vary in importance by stage of the life cycle.

More generally, even if none of the foregoing problems existed, I do not see how this type of analysis can demonstrate the value of a life-cycle variable. As is well known, income and spending patterns are influenced by many different variables, including education, occupation, age and family composition; plus the fact that a major determinant of expenditure allocation is the level of income. Only if these variables are taken into account in some multivariate framework is it possible to evaluate whether life cycle is a meaningful variable for the present purposes. In particular:

1. To establish the value of life cycle as an indicator of the allocation of spending patterns, one would have to show that a life-cycle variable is relevant after allowance has been made for income, education and occupation (or some combination of these in the form of socio-economic status).

2. To establish the value of a life cycle classification for explaining fluctuations in family income, one would have to show that this variable is relevant after taking into account education and occupation.

In both instances, one would also have to show that a life-cycle variable is more meaningful than, say, age or household composition considered separately.

The Paper by Landon and Locander

This paper is conceptually similar to that of the previous one. It seeks to establish the value of a life cycle variable for explaining variations in leisure behavior and attitudes toward such facilities. The approach is essentially the same as before, except that in this case the survey was carried out in Pima County, Arizona; the focus was on leisure time behavior and attitudes; and there is no information on response rates or on how the data were obtained. There is also no conceptual framework nor is there much review of the past literature on this topic, which I believe is extensive.

In view of the similarity of these two studies, the same sorts of questions can be leveled at this one as at the previous one. In particular:

1. There is no indication of the reliability of the data or of the survey.
2. It is not clear whether the particular life-cycle classification used is the most meaningful for the purposes.

From a more general point of view, the univariate analysis used in this study, with no attention to other relevant variables, provides no basis for inferring that this particular type of life-cycle classification is of any particular value. It could be, for example, that income or age, or a combination of the two, is a much better discriminator of behavior and attitudes toward recreation than a life-cycle variable. Until such alternatives are explored, the authors' contention of the value of a life-cycle variable remains unproven.

The Paper by Settle, Alreck and Belch

Methodologically, this study differs from the preceding two in that it includes five socio-economic and five demographic variables (one of which is life cycle) to investigate their relative effect in explaining participation in 100 different leisure time activities. Based on 1,000 chi-square tests that relate each of these variables to each of the activities, the authors conclude that the demographic variables are better predictors than the socio-economic variables, because education is the best single socio-economic determinant of leisure time behavior.

Unfortunately, as with the previous two papers, the validity of the differences is not supported by the analysis. About all that can be said from these data is that the demographic variables are more frequently associated with leisure time than the socio-economic variables, and that education is the variable that is most frequently significant. Such a result gives no indication of the relative importance of different variables. This is especially so since, here again, as in the preceding papers, no attempt was made to carry out any sort of multivariate analysis. Hence, we have no information about the relative importance of the different variables in explaining leisure time behavior, and there is no basis for doing so on the basis of the frequency of significance of explanatory variables to simple chi-square tests.

From a survey point of view, one could also raise a

number of questions on the procedures. Leaving the selection of a quota sample to be carried out by students in the field raises all sorts of possibilities of bias, both from a sampling and a response point of view, and none of these is given any consideration. Moreover, to have students offer to administer a personality test to respondents and provide them with interpretations is highly questionable from an ethical point of view; it is a procedure that should be strongly discouraged. (Incidentally, the assertion that mailing the interpretations of such tests to the respondents' homes provides validation of the data collection procedure has no basis, unless the respondents later were actually contacted.)

Finally, I do not understand what the authors mean by social status. Sometimes they seem to equate it with education; other times they seem to want to leave it to the judgment of the respondents; while still other times they suggest using analytical methods to obtain "candidate variables in determining social status, selecting those that perform the best". The concept itself is a highly ambiguous one and needs a great deal of additional work.

General Comments

As one may suspect from these comments, I am by no means convinced that the case for the value of life-cycle analysis has been established by these studies or, for that matter, by the studies carried out over twenty years ago in the conference volume edited by Lincoln Clark. This is not to say that I do not think that life-cycle analysis may not be a useful variable, but that the case for it has by no means been established one way or another.

To investigate if the life-cycle approach is useful in explaining consumer behavior would seem to involve two major steps, and preferably a third as well.

1. One step is to investigate alternative definitions of the family life cycle in relation to current living conditions. This means not only devising alternative definitions to take into account different forms of living, but also testing these definitions in terms of their ability to discriminate among different forms of consumer behavior. In this sense, it may well happen, for example, that a life cycle classification explaining one type of consumer behavior may not be the best for explaining some other type of consumer behavior.
2. Proceeding from these univariate tests, the best such classifications should be incorporated in a model of consumer behavior that includes other demographic and socio-economic variables. In a sense, this is the acid test, since if such a classification does not hold up when considered with other variables, it is of little analytical value. Such tests have rarely, if ever, been made, and it is only until such tests are carried out that we will have definitive information on the value of a life-cycle classification.

A third sort of test, one which would lend more generality to the preceding two steps, is to ascertain whether such results are valid across different cultures. Such cross-cultural comparisons are not easy, partly because comparable variables are not always available, and partly because errors in the data (to which practically nobody in the consumer behavior field pays any attention) may invalidate much of the results.

The latter point leads me to make the final comment that one thing badly needed in future papers of the type reported here is better reporting of the data collection process, and some evaluation of these data, an evaluation that is integrated with the substantive results. The

research on survey methodology has by now documented fully the substantial errors that can result from sloppy sampling procedures, failure to consider bias due to non-response, and failure to consider the effects of response errors, and of response variations. The least that consumer researchers can do is to be aware of these dangers, and to evaluate results they obtain in the light of the errors that are known to exist in the data.

DEMOGRAPHICS, SPENDING AND LEISURE:  
A CRITIQUE

Charles W. King, Purdue University

Abstract

The research papers presented in the session, Demographics, Spending and Leisure, extended historical demographic analysis in the areas of social class and family life cycle. The research, however, used theoretical and methodological approaches developed over ten to twenty years ago. Demographic analysis in consumer research demands new multidimensional demographic constructs that reflect the contemporary cultural milieu.

Introduction

The relationships between demographic variables and consumer behavior have been the focus of consumer research since the introduction of the analytical method in marketing. The relative effectiveness of demographic variables compared with other types of measures in explaining and predicting consumer behavior continues to be an unanswered question.

The three papers presented in the session, "Demographics, Spending and Leisure", relate traditional demographic variables to household expenditures by category and to leisure behavior. The papers represent a continuation of that historical demographic research tradition.

In reviewing the papers, significant issues of both a micro nature regarding the conclusions of the individual papers and of a macro nature regarding the entire concept of demographic analysis in consumer behavior have surfaced. Therefore, the critique of the papers has focused on:

- 1) the historical dimensionality of "demographic analysis" as reflected in the marketing literature;
- 2) a review and critique of the major conclusions and research issues presented in the papers;
- 3) the need for new demographic constructs in consumer research.

The Historical Dimensionality of "Demographic Analysis"

At the outset, a definition of terms is in order. What is "demographic analysis" as related to consumer behavior?

Demography is defined as "the statistical study of the characteristics of human populations, especially with reference to size and density, growth, distribution, migration, and vital statistics and the effect of all these on social and economic conditions" (Morris, 1971). Within the marketing context, demographic analysis involves "the study of market stratification of products or services in relation to such population characteristics as age, family size and type, sex, income, education and residence" (Ferber, et. al., 1964).

In the jargon of marketing, however, there is inconsistency in nomenclature regarding demographic and socioeconomic variables. Some researchers use the terms of demographic and socioeconomic variables interchangeably (Churchill, 1976). Other writers, however, draw clear differentiation between the two sets (Settle, Alreck, Belch, 1978). The pragmatic or theoretical implications

of the differentiation (or lack of differentiation) are not clear in the literature. That may be reflective of the current state of maturation of demographic analysis as a measurement construct.

Demographic and socioeconomic characteristics, however, are routinely used as identifiers of key target market segments. The demographic and socioeconomic characteristics of consumer market segments have been referred to as "states of being" because they identify attributes or profiles of people (Buzzell, et. al., 1969).

Related to other types of variables, demographic and socioeconomic characteristics are often "enabling" variables that make possible various forms of consumer buying behavior. Though "enabling" in character, demographic and socioeconomic characteristics are usually not sufficient to guarantee a particular buying behavior.

A consumer, for example, may have adequate disposable personal income to participate in a particular leisure activity, e.g., skiing. The income level may be "enabling" -- necessary -- to afford the sport expense, but not sufficient to guarantee personal commitment/involvement in that sport. "Interveneing variables" in the cultural milieu and in life style preferences may channel the consumer buying behavior into other alternative leisure activities.

In terms of demographic analysis, the first step in analysis is to relate a specific, unidimensional variable, age or marital status, to some aspect of consumer behavior, e.g., expenditures or leisure activity. As an extension of the unidimensional variable approach, demographic variables have been operationally "configured" or designed to be multidimensional. The new variables include combinations of standard demographic dimensions. Social class and family life cycle are two of the most widely used multidimensional demographic and socioeconomic variables measured.

Social class was first introduced as a socio-psychological concept in the United States over thirty years ago (Warner, 1949). Early pioneering work by the Chicago Tribune applied the notion in the marketing context in the 1950's (Martineau, 1957, 1958). Empirical definitions of social class typically involve the interactions of education, age, income and occasionally expenditure styles, e.g., expenditures on housing (Carmen, 1965).

The concept of family life style as a multidimensional demographic/socioeconomic variable was also developed in the marketing arena in the 1950's (Clark, 1955 and Wells and Gubar, 1966). The idea of family life cycle interrelates marital status, size of family, age of children, employment of family members, and residences of family members ("in the nest" or "out of the nest").

The basic postulate of family life cycle is that the process of family maturation over time can be organized around a series of discreet phases. Each phase has behavioral characteristics in terms of the social dynamics of the family unit. These dynamics ultimately impact on the family's consumer behavior in terms of both income availability and family needs and wants.

Interestingly, the concepts of social class and family life cycle as multidimensional demographic and socioeconomic variables are, in fact, first generation surrogate indicators of consumer life style segments. The validity of social class and family life cycle in consumer research is based on the crucial proposition that different social class segments or family life cycle stages are characterized by different behavioral patterns (life styles) that result in different consumer behavior. These variables preceded the development of life style and psychographics which emerged as a widely used research approach in the 1960's (Wells, 1974). Though some research has addressed this question, not enough interaction has developed between these closely related conceptual areas in the ensuing years (Reynolds and Wells, 1977).

#### Critique of the Research Contributions

##### An Overview

The research papers, in terms of a general overview evaluation, were conceptualized, empirically designed, and methodologically executed at the same level of research sophistication as the research traditions upon which they were built. Therein are both the strengths and weaknesses of the research contributions.

The papers universally represented extensions of the empirical data in the respective areas of social class and family life cycle research as applied to the analysis of consumer behavior. The papers did not, however, significantly extend the "state of the art" in the topic areas on either theoretical or methodological dimensions.

The Landon and Locander paper applied an established family life cycle paradigm with modifications to the analysis of leisure behavior research using a large sample survey research methodology. The conceptual approach, field measurement methodology, and analysis procedures of Arndt's research on family life cycle and household expenditure patterns used an established family life cycle measure with modifications. The field research survey was similar in overall design to statistical studies of consumer expenditures and construction/analysis of the Consumer Price Index routinely conducted by the United States Bureau of Labor over the same period (Consumer Expenditure Survey Series, 1972-1973). Likewise, the Settle/Alreck/Belch research used routine demographic and socioeconomic measures and analysis in focusing on social class as related to standard types of consumer leisure activity.

##### Significant Contributions of the Research

Each of the research studies made its own unique contribution to its respective research tradition and associated literature.

Landon/Locander established the validity of the family life cycle concept as a useful tool in understanding leisure and recreation behavior. The authors note that the family life cycle concept has potential for explaining leisure behavior because "it matches needs with groups of people".

Arndt delivered a valuable contribution in the cross cultural data on Norwegian consumer spending patterns and relating them to the family life cycle construct. Arndt concluded that "size and composition of household expenditures are systematically related to stage in family life cycle", and summarized that these findings have implications for welfare programs and consumer policies.

The Settle/Alreck/Belch research likewise related demographics and socioeconomic factors to leisure behavior. The researchers found that education was the best single socioeconomic determinant of leisure choices, contributing to the growing body of research on leisure behavior.

##### Significant Theoretical and Methodological Issues in the Research

In critiquing the research studies, significant theoretical and methodological issues were noted across the research projects. The research critiques have focused on four major dimensions for discussion: 1) literature review and theoretical rationale/"anchoring" of the research problems and designs; 2) construction of operational dependent variable measurements; 3) data collection and analytical procedures, and 4) interpretation of findings in terms of theory development or future research.

In the area of literature review and theoretical rationale/"anchoring" of the research problems and designs, all three papers could have presented more comprehensive reviews, research objectives and methodological rationale. The relevant theoretical frameworks, basic literature documentation and research issues were adequate and straight forward given the research objectives and the reporting medium.

The construction of the dependent variable measurements, however, lacked adaptive creativity and clarity of presentation. Admittedly, the measures were built upon historical research traditions of over twenty years and measurement comparability is essential for longitudinal comparison.

The issue is, however, that traditional dimensions of family life cycle and social class have potentially changed dramatically in the rapid socio/cultural transitions over the past decade. Those new dimensions were not explored or included in any of the operational measures used in the empirical research probing on family life style and social class as related to consumer behavior. Likewise, more detail on the specific measurements were needed in both the Landon/Locander and the Settle/Alreck/Belch papers.

On the dimension of data collection and analytical procedures, the Settle/Alreck/Belch data collection procedures involved an opportunity sample executed by 83 student interviewers conducting a class project. This raises the obvious problems of sample and interviewer methodology biases. Likewise, the use of a self-administered psychological test as an inducement for survey participation has serious implications in terms of bias, survey participant misinterpretation of the psychological test, and ethics of survey respondent manipulation.

On the analytical dimension, all of the reported analyses suffered from a lack of analytical depth and/or extension. The basic empirical data content was rich. All of the research, however, could have benefited from use of a higher order of analytical sophistication.

Presentation of detailed tests of significance would have been useful in both the Landon/Locander and the Arndt research. The Settle/Alreck/Belch research could have employed multivariate statistical analysis in addition to its univariate approach.

On the dimension of interpretation of research findings in terms of theory development and future research, all of the research papers emphasized data presentation, tabular display, and "numbers reporting". The papers needed more qualitative analysis and interpretation of the data in terms of the respective research traditions



and the "state of the art" of the topic areas upon which each project was built.

The Landon/Locander paper, for example, presented a very interesting data bank but concluded "FLC (family life cycle) offers a construct that could provide a rich independent variable to analyze present and anticipated recreational needs within the community. Likewise, FLC offers a means of segmenting markets into target groups whose research needs and subsequent behavior are relatively homogeneous." The researchers did not relate this data to the leisure research literature nor did they suggest issues for future research.

Arndt's expenditure pattern analysis could be viewed as merely a statistical report of Norwegian consumer expenditure patterns including a cursory qualitative summary of tabular results. By comparison, Arndt could have vastly enriched the research contribution by adding a cross cultural comparison of his data with similar, readily available published research conducted by the United States Bureau of Labor Statistics for the same year, 1972-1973. While the expenditure categories may not have been identically defined, there was enough comparability to support that comparative analysis and hypothesis generalization about cross cultural consumer expenditure dynamics.

By comparison, the Settle/Alreck/Belch research also failed to build any specific linkage between the research data and the research tradition with which the data related. On the other hand, these authors were very insightful in identifying and articulating the conceptual issues/challenges such as status inconsistency, the importance of various variables, e.g. education vs. income vs. occupational group, etc. in social class identification and/or determination. Additionally, the researchers presented a very valid discussion about the impact of cultural change on the concept and components of social class across the culture over time.

#### The Need for New Demographic Constructs in Consumer Research

Analysis of the historical dimensionality of "demographic analysis" and review of the research papers in this session generate a central concluding thesis. There is a dramatic need for development of new multidimensional demographic constructs in consumer research.

The concepts of social class and family life cycle as multidimensional demographic measures applicable to consumer market segmentation were theoretically conceptualized and empirically operationalized over twenty years ago. Very little methodological research has focused on the validity of the concepts as historically "configured" when applied in the contemporary era. The reported research used the concepts and their operational measures, with only minor modifications.

The cultural milieu has undergone traumatic social change over the past twenty years. That cultural change has impacted directly on the values of the social system and on the institution of the family.

Therefore, the central question is, are the historical measures of social class and family life cycle, in fact, applicable to the mass culture, circa 1978?

#### The Meaning of Social Class Under Social Change

The concept of social class, assuming Warner's original theoretical framework was valid, certainly deserves reassessment given the evolving values of the 1960's and 1970's. Social class is operationally defined within a specific societal context. Given the pluralistic character of the contemporary social milieu, can any

single definition of social class be meaningfully applied across the social system?

To suggest there is no social class dynamic in the contemporary society would be naive. To suggest that the social class dynamic can be portrayed by one macro aggregate social class measure that is over twenty years old may also be naive.

Looking to the traditional elements of social class, e.g. education, income, occupation, are these still the relevant inputs? What are their respective weights in determining social class...in what section of the social structure?

New social class measurements must be built upon assessment of the determinants of social class within particular societal situations. The current social class construct does not meet this challenge.

#### The Life Cycle and the Changing Family Structure

The concept of family life cycle may still be valid. Its operational measurement as reflected in the traditional nine stage continuum is not valid because it is not all inclusive. Over the past twenty years since the idea of family life cycle became popular, the institution of the family has experienced its own "future shock".

A variety of new family structures have evolved that are simply not identified in the current operational measurement of family life cycle. Because of the rising divorce rate, the one parent family has emerged as a structural entity that also flows through its own life cycle. The measurement of income, for example, becomes very complex within the divorced family community because of the confounding impact of alimony and child support payments on the disposable income of the receiver and on the payer. Likewise, the dynamics of child visitation cloud family membership measurement and significantly influence expenditure patterns.

Informal cohabitation is a type of family structure that has no convenient codification in the traditional family life cycle measures. That particular family arrangement, however, does influence consumer expenditure patterns particularly in housing and large ticket durable goods investments.

The working woman family has emerged over recent years as a unique type of family unit. Likewise, within the working woman families, an increasing share involve families with preschool children. In the context of family life cycle staging, the life style dynamics of this family segment are dramatically different from the classical, traditional profile of the "Full Nest, I (a)" stage as it has been defined in the operational measurements of family life cycle (Wells and Gubar, 1966). The implications for consumer behavior in the expenditure categories of child care and recreation are clear.

#### A Methodological Approach: The Convergence of Demographic Analysis and Life Style Research

The new demographic constructs can best be produced through a convergence of the two historical methodological traditions of demographic analysis and the maturing area of life style research. The multidimensional concepts of social class and family life cycle were both initially designed as surrogate life style descriptors prior to the development of the life style research tradition in the 1960-1970 era.

Social class and family life cycle, as demographic population descriptors, have been criticized, not because of their conceptual validity, but because of their operational measurement weaknesses. Nor are these

two dimensions the only potential multidimensional demographic measurements that might be informative to consumer researchers.

A complete review of demographic analysis is needed to identify relevant target consumer segments endemic to the contemporary culture. Life style dimensions must be explicitly combined with traditional unidimensional demographic variables such as age, income, education, marital status, employment, etc., to produce richer, more accurate and more relevant descriptions of the "status of being" of the evolving consumer market structure.

#### References

- Robert D. Buzzell, Donald F. Cox, and Rex V. Brown, Marketing Research and Information Systems: Text and Cases (New York: McGraw Hill, 1969)
- James A. Carman, The Application of Social Class in Market Segmentation (Berkeley: Institute of Business and Economic Research, University of California, 1965)
- Gilbert A. Churchill, Jr., Marketing Research (Hinsdale, Illinois: The Dryden Press, 1976).
- Lincoln H. Clark, ed. Consumer Behavior, Vol. II. The Life Cycle and Consumer Behavior (New York: New York University Press, 1955)
- Consumer Expenditure Survey Series: Interview Survey, 1972-1973, Report 455-3 (Bureau of Labor Statistics, Department of Labor)
- Robert Ferber, Donald F. Blankertz, Sidney Hollander, Jr., Marketing Research (The Ronald Press: New York, 1964)
- P. Martineau, Motivation in Advertising (New York: McGraw-Hill Book Company, 1957)
- P. Martineau, "Social Classes and Spending Behavior," Journal of Marketing, 23(October, 1958), 121-130.
- William Morris, ed. The American Heritage Dictionary of the English Language (New York: Houghton Mifflin Company, 1971)
- Fred D. Reynolds and William D. Wells, Consumer Behavior (New York: McGraw-Hill, 1977)
- W.L. Warner, Social Class in America - The Evaluation of Status (Chicago: Science Research Associates, Inc., 1949)
- William D. Wells, ed. Life Style and Psychographics (Chicago: The American Marketing Association, 1974)
- William D. Wells and George Gubar, "Life Cycle Concept in Marketing Research," Journal of Marketing Research, (1966), 355-363.

## WHAT DO YOU LEARN STANDING IN A SUPERMARKET AISLE?

K.W. Kendall, Simon Fraser University  
Ian Fenwick, Dalhousie University

### Abstract

During a five day period, over 2,300 shoppers were observed in supermarket aisles for 44 hours to determine the prevalence of label reading, the characteristics of the shoppers who read the labels, and the information format trade offs they would be willing to make. For the attention foods, 43% viewed the products for more than 8 seconds and had a mean viewing time of over 38 seconds. Discriminant analysis correctly classified between 75 and 90% of the defined shopper classifications. Grabbers and lookers showed marked differences in utility for both amount and type of information.

### The Provision of Consumer Information

Since long before the Fair Packaging and Labeling Act of 1966 in the U.S. and the revised Packaging and Labeling Act in Canada, consumer groups had demanded further information be available on food products so consumers could make better food choices. Better information has come in the form of unit pricing, ingredient information lists, open dating and the presentation of nutrient content, etc. The provision of this information was heralded as the solution to the information gap for consumers.

However, some early work on information load in a laboratory setting (Jacoby, Speller and Kohn, 1974a, 1974b), suggested that more information produced "dysfunctional consequences" for consumers. A key issue involved the definition of information, e.g. more brands as more information or more information on each product. Several researchers have pointed to this problem (Russo, 1974; Summers, 1974; Wilkie, 1974). In fact, Russo (1974) suggests that more information is better since evoked set size is usually found to be quite small (Howard, 1977).

Friedman (1972), reviewing the empirical literature on consumer use of information aids in supermarkets and on food labels, found most studies to have identified relatively few regular users of the three aids most extensively researched: unit pricing, open dating and nutrient labeling. Friedman's conclusions are somewhat misleading. At the time of the review a lot of the aids were relatively new. Furthermore, the dependent variables were structured in a normative sense of "best choice". Friedman defines regular use by whether or not consumers could correctly identify or reason through specific names of the new aids. More recently, the same author (Friedman, 1977) has reported little reason to alter those largely negative findings.

Reviewing similar work, Ross (1974) has suggested that a hierarchical relationship exists. Many people are aware of the information aids but few report using them. Day (1976) and Jacoby, et al (1977) echo this hypothesis of a hierarchy of effects. Day points to the generally large number of consumers who appear aware of the information in supermarkets and on the labels of food but the small number of people who claim to use this information. Interestingly, the hierarchy has not been tested in the stores. Instead, empirical studies rely on survey self-reports or laboratory situations with highly motivated subjects. There is little actual verification in the field.

The low usage of label and supermarket information aids is frequently blamed on the consumers' lack of comprehension of such information as noted by Daly (1976) and Jacoby, et al (1977). This has been true of open dating methods (Taylor, 1976), ingredient names (Warland and Herrmann, 1971), unit pricing measures (King and Gideon, 1971) and nutrient information (FDA, 1973, 1975; Kendall 1977).

As a result, the most recent emphasis has been on methods of presenting information and information processing formats (Kendall, 1977; Scammon, 1977; Bettman and Kakkar, 1977; Russo, Krieser, and Miyashita, 1975). It should be emphasized that many public policy decisions are presumably based on survey self-report evidence (FDA, 1973, 1975) and assume the hierarchy outlined by Ross (1974) and Day (1976). However, Day himself does not pretend to present an extensive literature review. And it is not at all clear that the evidence does support the hierarchy hypothesis.

A more extensive literature review on consumers' importance ratings for information aids shows that price information tends to be most important followed by brand name and then nutrition values in the aggregate data (FDA, 1975, 1973; CRI/FDA, 1972; Lenahan, et. al., 1972; Darden and French, 1971). These studies also show a trend more recently of evaluating nutrition cues more heavily than price users.

Even more emphatic are the reports on use, or probability of use, of information items on the labels of food products. Two points are of extreme interest. First, the more recent the study, the higher the reported use of different information aids (Jacoby, et. al., 1976; FDA, 1975; Babcock and Murphy, 1973; FDA, 1973; Lichtenstein, 1972; Lenahan, et al. 1972; CRI/FDA, 1972). For example, the CRI/FDA (1972) study found only 24% of the respondents reported checking the ingredients list. By 1975, this figure had more than doubled (FDA, 1975).<sup>1</sup> Second, claimed usage rates are not as much below the importance ratings as hierarchy models would suggest.

Furthermore, the Redbook Nutrition Study (1974) found almost half their sample claiming regularly to read food content information, and a further 45% claiming "occasionally" to read such information. For new products claimed readership was rather higher, 60% regularly reading, 36% occasionally reading. More recently, an even higher figure is reported, 66% regular reading (Redbook Nutrition Study, 1976).

However, there appears to be little evidence from the market place to determine how many consumers look at the information on products, the kind of people who actually look at labels or their preferences for label information formats in the store. Best and McCullough (1977) have looked at the consumer utility of label information with conjoint measurement, but did not tie their preference measures to any strong behaviour measures. Furthermore, Day (1976) suggests that there is a pressing need for field research rather than laboratory studies in this area.

<sup>1</sup> For further details on these figures, see Kendall and Fenwick, Simon Fraser Univ. Working Paper, 1978.

The studies reported here were designed to accomplish three tasks. First, the actual number of people who look at labels and the amount of time they spend looking at food products in a regular supermarket setting was determined. The emphasis on the time construct has recently been noted in the literature (Jacoby, Szybillo and Berning, 1976). Time estimates from actual shopping observations should also be helpful in verifying laboratory studies on information processing of information formats.

The second purpose of the studies was to determine the characteristics and situations of those consumers who look at food products and those that do not look at food products. Finally, the studies were developed to measure the utility of different information formats for a new convenience food product for the different shopper groups.

#### Methodology

The data analyzed here were collected in two separate studies. Both were run in Halifax, N.S. where the only required label information is a list of ingredients in descending order of proportion (Health Protection and Food Laws, 1970). The studies were conducted in two of the larger supermarkets in the metropolitan area, one from each of the major chains. Although basic procedures were identical in both studies, the products examined differed and there were minor differences in the questionnaires used. In what follows the particular products studied will be referred to as the attention foods or products.

In both studies two researchers, dressed as supermarket personnel, were stationed in an aisle and recorded the time spent looking at the attention foods by every shopper entering that aisle. A sample of the shoppers who had passed through the aisle was then interviewed by other researchers in another part of the store. Stratified random sampling was used, strata being defined by the time spent looking at the attention foods.

The first study, referred to here as AISLE 1, concentrated on rice and pasta products. These foods were chosen because previous studies had reported that particular attention was given to their labels (see for example FDA, 1975). Observations were taken throughout shopping hours on a Monday and a Thursday. A total of 1,328 shoppers was timed and 141 were interviewed during the two day period. For the second study, AISLE 2, the attention products were canned meat/fish and powdered (dehydrated) soup. The canned meat/fish products were chosen because observations during the earlier study was suggested that shoppers paid particular attention to those foods. Powdered soup was selected as representing a new convenience food, heavily prompted at the time by mailed samples, and marketed at least partly on a health basis. Observations were taken throughout shopping hours on a Wednesday, Thursday, and Friday. For the second study, a total of 1,053 shoppers was timed and 147 were interviewed over the three day period

The personal interview questionnaire on the AISLE 2 study was shorter than that for AISLE 1 to allow respondents time to perform a label ranking task. For this respondents were shown nine labels for a new food product, each presenting the same product information but in different formats. The subjects were asked to place the labels "in order of their helpfulness to the subject for making a decision to buy a new food product." The nine labels formed an orthogonal design embodying three levels for each of four types of information: brand name, ingredient information, nutrition information and open dating information. The levels used for each type of information are detailed in Table I. This fractional factorial design affords extreme economy in treatments (full factorial design would involve 81 product labels)

allowing the ranking task to be kept within individual capabilities. Clearly, the cost of such parsimony is the assumption of no interaction effects; however, previous studies suggest such interactions are rare (see for example Green and Wind, 1975; Green, 1974). An artist's sketch of one of the labels used appears in Figure I.

#### The Time Construct and Shopper Classifications

The major dependent variable in the studies reported here is the time each shopper spent examining the attention products. Jacoby, Szybillo and Berning (1976) have reviewed the literature on time and consumer behaviour. As they point out very convincingly, no major conceptual treatment nor systematic empirical effort has yet been focused on this subject. Our studies are an attempt to look at the time construct in relation to consumer behaviour unobtrusively. There appears to be no such studies reported in the literature; if there is one it is the exception.

TABLE I: DESIGN FOR CONJOINT ANALYSIS

F1	BRAND NAME	1. NO BRAND NAME (EBERT, 1977) 2. NATIONAL BRAND NAME (MFG) 3. PRIVATE BRAND NAME (STORE)
F2	NUTRITION INFORMATION	1. SCOREBOARD METHOD (JACOBSON, 1973) 2. FDA/FTC METHOD 3. FOOD EQUIVALENT METHOD (BABCOCK, 1971)
F3	INGREDIENT INFORMATION	1. REGULAR SET OF NAMES (REQUIRED INFORMATION) (RN) 2. RN + MEANING OF THE INGREDIENT NAMES (M) 3. RN + M + WARNING/COMMENT
F4	OPEN DATING INFORMATION	1. MANUFACTURERS' DATE, e.g. 1 Sept 1977 2. SELL BEFORE DATE, e.g. 1 May 1978 3. CONSUME BY DATE (BEST BEFORE), e.g. 1 Feb 1978


FRACTIONAL FACTORIAL DESIGN --  $3^4$  ORTHOGONAL DESIGN<sub>2</sub>

1	1	1	1
1	2	2	3
1	3	3	2
2	1	2	2
2	2	3	1
2	3	1	3
3	1	3	3
3	2	1	2
3	3	2	1

Nine cells representing the 81 cell complete design. The subjects were asked to rank the nine labels. The labels were presented to each subject in random order

2. Addelman (1962) presents rules for the construction of these designs. Basically for any  $(s^n - 1)/(s - 1)$  factors each at  $s$  levels, an orthogonal design can be constructed using only  $s^n$  treatments.

FIGURE I: ILLUSTRATIVE LABEL

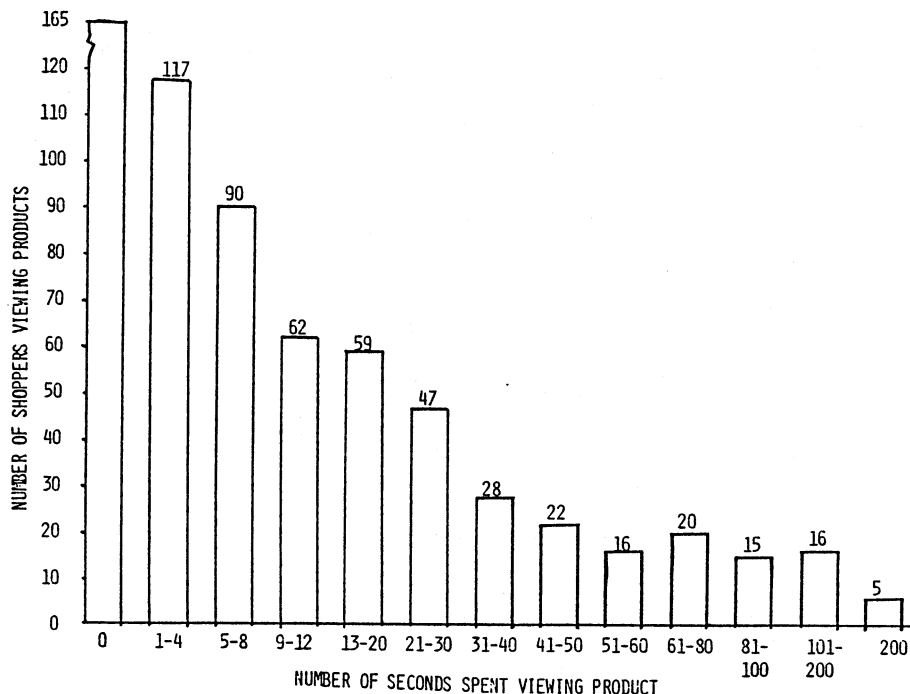
<p><b>NEW BACON SUBSTITUTE</b></p> <p>454 g</p>  <p><b>DIRECTIONS</b> Cover bottom of skillet with oil and cook over medium heat 6 to 8 minutes. Do not overcook.</p> <p><b>KEEP REFRIDGERATED</b></p> <p>Manufacturer Date: Sept. 1/77</p>	<p><b>INGREDIENTS</b> Pork, Bacon, Sugar, Sodium Phosphate, Water, Flour, Skim Milk Powder, Salt, Beef By Product, Spice, Sodiumerythor Monosodium Glutamate, Sodium Nitrite, Corn Oil</p> <p><b>NUTRITION SCOREBOARD</b> Per Serving Serving Size = 4 Strips Servings Per Container = 4.5</p> <p><b>SCORE -155</b> Score Range = 0-280</p>
---	---

The time construct is emphasized here both because little is known about how consumers allocate their time in the supermarket, and because the time spent viewing products is one operational measure of the use consumers make of labels; extremely short viewing times suggest virtually no use of label information whereas extended product viewing implies considerable processing of label information and product comparisons. This is further supported by Gatewood and Perloff (1973) who have suggested that speed in processing supermarket and label aids would seem to be a desirable end product of consumer information systems.

Figure 2 shows the frequency distribution of time spent examining the attention products for all shoppers entering the relevant aisle throughout both studies. The distribution is heavily skewed towards short product viewing times, but there is a sizeable tail with sixteen individuals actually examining the products for more than two minutes. Given that the interest is in product viewing time as a measure of the use and interest in label information and that the distribution in Figure 2 is so heavily skewed, it is preferable to treat it as a classificatory rather than a continuous variable.

The obvious dichotomy was between "grabbers" who simply snatched the product from the shelf with virtually no discernible product viewing time, and "lookers" who observed the products for at least one second. In view of the wide range of viewing times by "lookers" (from one second to 5 1/2 minutes) this is a very crude classification. Consequently, for parts of the analysis this group was subdivided into "scanners" who observed the products under study for eight seconds or less and "label readers" who considered the products for more than eight seconds. The eight second cut-off aimed to distinguish between shoppers employing a routinized response and those in a limited or an extensive problem solving stage of decision making (see Howard, 1977).

FIGURE 2: FREQUENCY DISTRIBUTION OF PRODUCT VIEWING TIMES



As discussed previously there is little prior research on the role of time in consumer behaviour and the 8 second cut-off time demands some explanation. In the design of this study estimates of reading times were obtained using a group of undergraduate students. (For details of this methodology see Kendall and Fenwick, 1978.) The results suggested an average scanning time of about eight seconds.

This crude estimation does not appear to be too inconsistent with the literature. On an average package of food, a reader can identify about seven chunks of information: price, brand name, contents, ingredients, directions, nutritive composition (or where to get it), weight, manufacturer's name and address, and a few aesthetic phrases about the glories of the product. Peterson and Peterson (1959) and Murdock (1961) suggest that percentage of correct recall tends to level off at about eight seconds for three consonant trigrams. This would be equivalent to three chunks of information which the average consumer might choose to look at on the label of a food package. The above time referents are based on a retention task. Most food shoppers are not trying to retain the information; they are trying to make a decision with the information which would entail some retrieval time.

Recent work with chronometric analysis in consumer behaviour (Gardner, Mitchell, and Russo, 1977) for low/high involvement products showed that response times for brand and attention evaluation statements averaged just under four seconds. Furthermore, Johnson and Russo (1977) have shown that mean recall times for a brand probe and an attribute probe for cooking oil was just over 12 seconds. However, the latter again refers to recall time and most shoppers are not concerned with this task. One could argue that a mean time might be 8 seconds merely by combining the above two studies. But it would be more realistic to take the first study and double the four seconds since the shopping is not a forced choice situation and shoppers do not feel that they are being tested. Consequently, the cut-off between scanners and readers of 8 seconds does receive

some justification, both in the literature and from pilot empirical work. Clearly this is an important area for further field research.

### Results and Discussion

#### Prevalence of Product Reading

During the five days covered by the studies, 2,381 shoppers passed through the attention aisles. Summary statistics for the various days, products and shopper categories appear in Table 2. Three points are of particular interest.

First, comparing results across product groups, rice and pasta buyers are four times more likely to simply grab the product than are canned meat/fish and soup buyers. This finding conflicts completely with the FDA (1975) study which reported grain products (bread, rice and pasta) as the most "label read" food products. The FDA study, relied, of course, on claimed label usage in a personal interview situation. It would appear that such claims may not be a reliable indication of behaviour and should have little input to decision making.

Second, over both studies (i.e., including the rice and pasta products), only 25% of buyers simply snatched the product from the shelf. The other 75% showed at least some interest in gathering product information. This contrasts with the self-report surveys which (with the exception of the Redbook Studies 1974, 1976) generally report around 30-50% claimed product information usage. Furthermore, of those showing some interest, over half spent more than 8 seconds with the products, giving a mean product viewing time of 38 seconds. Although no specific information was gathered on the particular product/package/label attributes that warranted this attention time, the extent of exposure suggests much more information processing, or product comparisons, than is normally assumed (for example see Friedman, 1977, reviewed earlier).

Third, looking in more detail at the difference in reading time between products, canned meat/fish recorded the longest mean time (42 seconds) with the newer product, dehydrated soup, having the shortest (34 seconds) -- a difference significant at the .025 level. The label reading literature would suggest the reverse effect. Viewing times are hypothesized as being shorter for the "older" more established products (see for example Howard, 1977; Buck and Jacoby, 1974; Ross, 1974). Interestingly all these studies have concentrated either on pure theory (e.g., Howard, 1977) or else have been confined to forced-choice laboratory situations (e.g., Buck and Jacoby, 1974). Clearly there are more complex factors at work than has been popularly supposed.

TABLE 3: STANDARDIZED DISCRIMINANT COEFFICIENTS

Variable	2-Way Analyses		3-Way Analysis	
	Aisle 1	Aisle 2	Function 1	Function 2
Day of the week	.61	-.73	.82	.21
Time of day	.88	-.55	.57	.03
Age of shopper	.37	.31	.82	.01
Marital status	.10	.19	.21	.20
Shopping party size	-.54	.49	-.64	-.21
Use of shopping list	-.28	-.14	-.54	-.15
<b>Additional Variables</b>				
Satisfaction with current labelling information	-.23	-	-.31	-.13
Support of mandatory Federal label regulation	.38	-	.22	-.49
Degree of menu-planning	1.04	-	1.00	-.39
Years of full-time education	-	.40	-	-

TABLE 2: SUMMARY STATISTICS OF PRODUCT VIEWING TIME

DAY	PRODUCT	#'s WALK THRU	#'s GRABBERS 1	#'s SCANNERS 2	#'s LABEL READERS 3	TIME VIEWING PRODUCT		TOTAL 5's	DIFFERENCES IN READING TIME FOR LABEL READERS
						SCANNERS	LABEL READERS		
MON	Rice	233	20	15	8	4"	39"	276	t = .970, p < .25 DF = 2
	Spaghetti	103	18	10	7	4"	27"	138	
THUR	Rice	362	26	24	15	5"	39"	427	t = .884, p < .25 DF = 2
	Spaghetti	351	57	50	39	5"	34"	487	
WED	Canned Meat and Fish	63	4	16	41	4"	42"	124	t = 1.66, p < .10 DF = 3
	Powdered Soup	116	4	12	18	4"	34"	150	
THUR	Canned Meat & Fish	72	13	34	79	4"	56"	198	t = 2.88, p < .025 DF = 7
	Powdered Soup	144	4	28	23	5"	36"	199	
FRI	Canned Meat	163	10	11	16	4"	41"	200	t = 1.07, p < .25 DF = 2
	Canned Fish	112	9	20	41	4"	27"	182	
TOTAL		1719	165	210	287	4.3"	38"	2381	
% of all shoppers		72%	7%	9%	12%			100%	
% of shoppers interested in prod.			25%	32%	43%	n = 662			

1. Less than one second viewing product 2. One to eight seconds viewing product  
3. More than eight seconds viewing product

## Who Looks At Labels?

In answering this question the primary distinction made was between "grabbers, who spent negligible time with the product, and the rest (i.e., scanners and label readers). Non-buyers were excluded completely from the analysis. Two-way discriminant analyses were performed separately for the AISLE 1 and AISLE 2 data.

First, restricting the analysis as far as possible to variables common to both data sets, a stepwise analysis was performed for AISLE 1. This found six variables (chosen from a total of 15) which could correctly classify 75% of the sample.<sup>3</sup> These six variables were: days of the week, time of the day, age of the shopper, marital status, shopping party size and use of a shopping list.

People who spent some time looking at the products tended to be older, married, without a shopping list, in smaller shopping parties, and to be shopping later in the week and later in the day than grabbers (see coefficients in Table 3). Interestingly, self-reported in-store label reading behaviour did not discriminate well. If anything, grabbers tended to claim more label-reading than did lookers! Thus, self-reported label reading was found to be an unreliable indicator of true behaviour, at least for the rice and pasta products studied in AISLE 1.

Widening the predictor set to include three variables not measured in the AISLE 2 study - satisfaction with current label information, extent of support for mandatory federal label regulation, and the degree of menu planning - improved the correct classification rate to 90%. People who looked at labels tended to plan menus further ahead, to support mandatory federal regulation of label information, and to be dissatisfied with the current label information offered (see Table 3). Although the latter two effects are intuitively appealing, it is less obvious why menu planners tend to examine labels when those with shopping lists do not. One possible explanation is that menu planners spend the time with labels to determine if the product fits into the menu scheme. This would indicate the consumer uses either direction or recipe information, should it be available. The shopping list person on the other hand simply wants the specific product, thus precluding the need for information in the store.

Having successfully distinguished grabbers from the product lookers, a three-way discrimination was attempted. Those who looked at products were divided into scanners (looking for 1-8 seconds) and readers (looking for more than 8 seconds). The same variables described previously produced a 74% correct classification rate using two discriminant functions (see Table 3, column 3). Interestingly, the first discriminant function was almost identical to that estimated in the two-way analysis, and label-readers were simply like the lookers discussed above, but more so! On the second discriminant function however, grabbers and readers tended to have similar scores, with the scanners well apart from both. It seems then that in some respects label readers have more in common with grabbers than scanners. One possible

<sup>3</sup> The dangers of evaluating a discriminant function by correct classification rate are well known (for example, see Morrison, 1969). In this case validations using the jack-knife method (Tukey, 1958; Fenwick, 1978) are currently being performed. Furthermore, in no group was the correct classification rate more than 3% below the average reported.

explanation for this would be that grabbers are currently processing little information, but if alerted, say by a health scare, search for product information, i.e., become readers. Scanners on the other hand are already gathering and processing the available information in under 8 seconds.

Applying the original six variables to the AISLE 2 data set produced an overall average 70% correct classification rate. Although using the same variables, unfortunately the AISLE 2 discriminant function showed some sign changes (see Table 5). For the AISLE 2 products (canned meat/fish and powdered soup) people who look at the products, although still older and married, tend to have a shopping list and to shop earlier in the week, earlier in the day and to be in larger shopping parties. Part of the explanation for this is due to the differences in products and the different store chain used in the second study.

As always when using search procedures, invalidated results must be treated with caution. It is impossible to say at this stage whether these sign changes show unstable discriminant functions or identify true differences in the characteristics of label readers for different products and different stores. Research is proceeding on this point using the jack-knife method (Tukey, 1958) to test the stability and significance of the discriminant results.

Substitution of years of full time education for age improved the correct classification to 75%. Those looking at labels were better educated than the grabbers as the literature would suggest (Daly, 1976). A three-way discrimination between grabbers, scanners and readers could not be successfully performed for AISLE 2, correct classification being well below 50%.

In summary, the question of who looks at labels is not easily answered. Although the same sets of variables identify product lookers in both studies, the direction of influence of these variables differs.

## Shoppers' Trade-Offs For Label Formats

In the final part of the AISLE 2 study shoppers were asked to rank order 9 labels for a new food product (see Figure 1) according to their helpfulness in making a buying decision. These 9 labels formed a fractional factorial representation of a 3<sup>4</sup> design as noted in Table 1. Interest centered on a new food product since regulatory authorities are particularly interested in assisting consumers with new product choices (U.S., 1974). Preference rankings were analyzed for a group of lookers using the MONANOVA program.<sup>4</sup>

Figure 3 shows the part-worth utility functions for each of these groups. Almost all the functions show marked differences between grabbers and lookers. Figure 3a shows the part worths for different forms of nutrition information. Results are consistent with the literature although the Food Equivalent Method performed less well than would be expected from other results (Babcock and Murphy, 1973). Lookers tended to value the more extensive FDA/FTC Method, while the grabbers preferred the singular representation of the Jacobson (1973) Scoreboard Method. Similar results were shown for these methods in media presentations (Kendall, 1977; Kendall and Krane, 1978).

<sup>4</sup>The preference rankings of individuals within these groups were treated as replications, a single set of trade-offs being estimated for the whole group. The alternative procedure of estimating trade-offs for each individual and then clustering the results (e.g., Best & McCullough, 1977) makes interpersonal comparisons of individuals' trade-offs which are strictly inadmissible and produces less well constrained solutions.

Figure 3b notes the trade-offs for different ingredient information. It is not surprising that both groups have little liking for the present method of presenting ingredient information since it is not easily understood (see G.A.O., 1975). The grabbers again prefer the succinct warning/comment statement since presumably it takes less time and is more convenient. The lookers have the highest utility for the explanation of the ingredients and appear to think that a warning/comment does not add to their information handling capacities.

Figure 3c shows the part worths for open dating formats. The "sell before date" appears to offer the highest utility for both shopper groups. However, the difference in utilities over all three formats is fairly small indicating the low importance of this piece of information to both groups.

Finally Figure 3d shows shoppers' reactions to brand name information. For the typical grabber private brand name has highest utility and the no brand name label introduced by some major retailers (e.g., Carrefour, Jewell and Dominion, see Ebert, 1977) is least preferred. In contrast product lookers actually prefer no brand products. While it is understandable that product lookers do not need the additional information offered by brand name - they can acquire the necessary information from other items on the product label - it is difficult to see why they should actually prefer the no brand name as an information aid. Presumably other characteristics of no brand products (e.g., low price) led to its higher preference ranking.

FIGURE 3: PART WORTH UTILITY FUNCTIONS

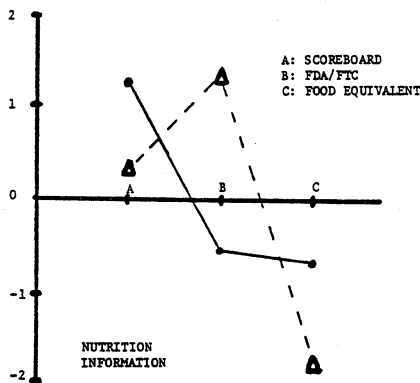


Figure 3a: Nutrition Information

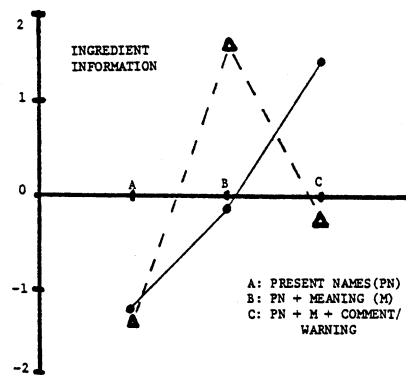


Figure 3b: Ingredient Information

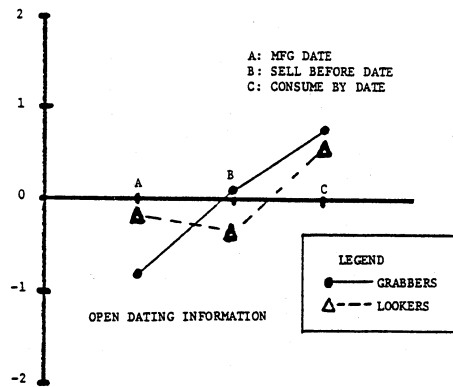


Figure 3c: Open Dating Information

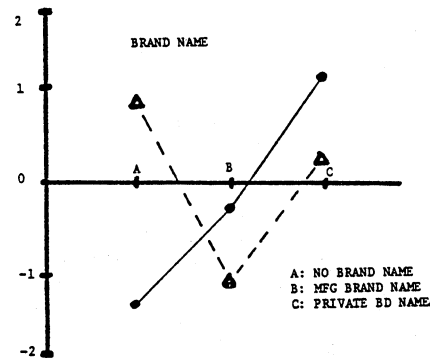


Figure 3d: Brand Name

Examining the range of utility covered by each information aid gives an indication of the importance of that kind of information to shoppers (Green and Wind, 1975). Table 4 summarizes the results. For the average grabber, brand name and ingredient information constitute over 60% of the utility range, whereas for lookers nutrition and ingredient information account for almost 70% of the total range in utility.

Of particular interest is the difference in the order of importance of information items between the two groups. For the average grabber brand name is the most important information source, followed by ingredient information. The average looker, however, is most concerned with nutrition information, although ingredients are a close second. The importance of nutrition information when buying a new food product is most interesting as Canada, unlike the U.S.A., has no mandatory legislation on nutrition information. Ottawa, and Consumer and Corporate Affairs, has been reticent about such legislation and little empirical research has been reported (see Liefeld and Bond, 1974; Liefeld, et al., 1975).

The only other reported work on label information trade-offs appears to be that by Best and McCullough (1977). That study collected its data by in-home interviews and could not, therefore, relate results to strong behavioural measures. Furthermore, most of the information formats tested in that paper are neither in use nor even proposed. Finally, the Best-McCullough study sums utilities across individuals, an operation which is not meaningful with MONANOVA output.

Conclusions

First, this study found self-reports of label reading



to be misleading. Food products for which other studies had found high self-reported label reading turned out not to be read frequently. Individuals who reported label reading in this study spent, if anything, less time with the product than did those who claimed little reading. For public policy purposes any studies based on self-reported label reading should be treated with caution.

TABLE 4: SUMMARY OF UTILITY RANGES FOR INFORMATION AIDS

Grabbers

	<u>Difference</u>	<u>Importance</u>	<u>Rank</u>
Brand Name 1.634 - (-1.326) =	2.96	32%	1
Nutrition Formats 1.224 - (-1.668) =	1.89	21%	3
Ingredient Info. 1.468 - (-1.339) =	2.80	30%	2
Open Dating .766 - (-.817) =	1.58	17%	4
<u>Total Difference</u>	<u>9.23 = 100%</u>		

Lookers

	<u>Difference</u>	<u>Importance</u>	<u>Rank</u>
Brand Name .843 - (-1.075) =	1.92	21%	3
Nutrition Formats 1.375 - (-1.738) =	3.11	35%	1
Ingredient Info. 1.631 - (-1.374) =	3.0	34%	2
Open Dating .512 - (-.387) =	.90	10%	4
<u>Total Difference</u>	<u>8.93 = 100%</u>		

The Ideal "grabber" label

1. Private Brand Name
2. Ingredients and Meaning and Warning/comment
3. Scoreboard nutrition method
4. Best before date

The Ideal "looker" label

1. FTC/FDA Nutrition Format
2. Ingredients and Meaning
3. No Brand Name
4. Best before date

Second, a large majority of shoppers spent enough time with the products as to imply considerable label reading. Only 25% simply snatched the product from the shelf. The rest had time enough to process at least some label information and some had time to process large amounts of information. However, product class strongly influenced consumers' propensity to view the product.

Third, although using a limited number of descriptor variables, this study suggests that it is not as much the type of person but the situational variables which determine label reading. Day of the week and time of day play a large role in determining readership. Degree of menu planning and possession of a shopping list also determined readership. The literature takes little account of these variables and further research should look into such explanations.

Finally, ingredient and nutrient data are valued by consumers. Even those simply grabbing the product from the shelf admit the usefulness of ingredient information. For the majority of shoppers nutrition and ingredient information dominate their label preferences for new products. However, a public policy decision has to be made on the target market for such information. It is not enough to say that there are different criteria for different groups. If information is to appeal to grabbers, who could be in need of more guidance, ingredient information should contain simple warnings, and nutrition data should be in the scoreboard form. On the other hand, for lookers, who are more likely to process the information, more detailed formats are appropriate. There is no real evidence that this issue has yet been confronted by public authorities.

References

S. Addelman. Symmetrical and asymmetrical fractional factorial plans. Technometrics, 1962, 4, 21-58.

Babcock, M.J. & Murphy, M.M. Two nutritional labeling systems. Journal of the American Dietetic Association, 1973, 62, 155-161.

R. Best, & J. McCullough. Evaluation of food labeling policies through measurement of consumer utility. In Hunt, K. (Ed.). Advances in Consumer Research, 1977, 5, 213-219.

M. Buck & J. Jacoby. Time spent acquiring product information as a function of type of innovation and information load. Purdue Papers in Consumer Psychology, No.136, Purdue University, 1974.

R.W. Chestnut. The Expenditure of Time in the Acquisition of Package Information. Unpublished Masters Thesis Purdue University, 1975.

CRI/FDA. R.C. Stokes & R. Haddock. Interim report of the first two phases of the CRI/FDA nutritional labeling research program. Washington, D.C. Consumer Research Institute, 1972.

P.A. Daly. The response of consumers to nutrition labeling. Journal of Consumer Affairs, 1976, 10, 170-178.

R.W. Darden & W.A. French. Product label information: Dimensions of meaning and evaluation of importance. Combined Proceedings of the Spring and Fall Conferences of the AMA, 1971, 647-652.

G.S. Day. Assessing the effects of information disclosure requirements. Journal of Marketing, 1976, 40, 42-52.

L.K. Ebert. No-Brand lines bow at Jewel Stores. Advertising Age, 31 October 1977, 2.

FDA, U.S. DHEW. Food and Nutrition: Knowledge and Beliefs, Report I, 1973-4 (RAC#3697). Washington, D.C. 1973.

I. Fenwick. A user's guides to conjoint analysis in marketing. European Journal of Marketing, 1978, 12, forthcoming.

M. Friedman. Consumer use of Informational aids in supermarkets. The Journal of Consumer Affairs, 1977, 11, 78-89.

M. Friedman. Consumer responses to unit pricing, open dating, and nutrient labeling. Proceedings of the 3rd Annual Conference of the Association for Consumer Research, 1972, 3, 361-369.

GAO, Comptroller of the U.S. Food Labeling: Goals, Shortcomings and Proposed Changes, Washington, D.C.1975.

M.P. Gardner, A.A. Mitchell & J.E. Russo. Chronometric analysis: An introduction and application to low involvement perception of advertisements. In Hunt, K. (Ed.). Advances in Consumer Research, 1977, 5, 581-589.

R.D. Gatewood & R.B. Perloff. An experimental investigation of three methods of providing weight and price information to consumers. Journal of Applied Psychology 1973, 57, 81-85.

P.E. Green. On the design of choice experiments involving multifactor alternatives. Journal of Consumer Research, 1974, 1, 61-68.

P.E. Green & Y. Wind. New way to measure consumers' judgments. Harvard Business Review, 1975, 53, 107-118.

J.A. Howard. Consumer Behavior: Application of Theory. New York: McGraw Hill, 1977.

M. Jacobson. Nutrition Scoreboard: Your Guide to Better Eating. Center for Science in the Public Interest, Washington, D.C. 1973.

- J. Jacoby, R.W. Chestnut & W. Silberman. Consumer use and comprehension of nutrition information. Journal of Consumer Research, 1977, 4, 119-28.
- J. Jacoby, D.E. Speller & C.A. Kohn. Brand choice behavior as a function of information load. Journal of Marketing Research, 1974a, 11, 63-69.
- J. Jacoby, D.E. Speller & C.A. Kohn. Brand choice as a function of information load: replication and extension. Journal of Consumer Research, 1974b, 1, 33-42.
- J. Jacoby, G.S. Szybillo & C.K. Berning. Time and consumer behavior: An interdisciplinary overview. Journal of Consumer Research, 1976, 2, 320-339.
- E.J. Johnson & J.E. Russo. The organization of product information in memory identified by recall times. In Hunt, K. (Ed). Advances in Consumer Research, 1977, 5, 79-86.
- K.W. Kendall. An experimental study to investigate methods of presenting nutrient values in the media. Unpublished Doctoral Dissertation, The University of Iowa, 1977.
- K.W. Kendall & R. Krane. Television, nutrition and product choice. Simon Fraser University Working Paper, 1978.
- K.W. Kendall & I. Fenwick. What do you learn standing in a supermarket aisle?, Simon Fraser University Working Paper, 1978.
- J. King & J. Gideon. Unit pricing for new grocery system Report prepared for Delchamps, Inc., Mobile, Alabama, 22 December 1971.
- R.J. Lenahan, J.A. Thomas, D.L. Call, & D.I. Padberg. Consumer reaction to nutritional information on food product labels. Search Agriculture, 1972, 2, 37-53.
- J.P. Liefeld & E. Bond. Consumer information schemes - Do we need them, do we want them, will we use them. The Canadian Consumer, June 1974, 16-18.
- J.P. Liefeld, E.A. Gullett, M. Sanderson, T.A. Watts & J.H. Sabry. Attitudes of a sample of consumers towards nutritional labeling. Canadian Institute Food Science Technology Journal, 1975, 8, 46-50.
- D.G. Morrison. On the interpretation of discriminant analysis. Journal of Marketing Research, 1969, 6, 156-163.
- B.B. Murdock. The retention of individual items. Journal of Experimental Psychology, 1961, 62, 816-825.
- L.R. Peterson & M.J. Peterson. Short term retention of individual terms. Journal of Experimental Psychology, 1959, 58, 193-198.
- Redbook Nutrition Study. Conducted for Redbook Magazine by the Gallup Organization Inc., 53 Bank Street, Princeton, N.J. May 1976.
- Redbook Readership Survey. How do you feel about food, what do you know about food? Conducted by the Editorial and Research Departments of Redbook Magazine, 1974.
- I. Ross. Applications of consumer information to public policy decisions. In Sheth, J.N. & Wright, P.E. (Ed). Marketing Analysis for Social Problems. Illinois: University of Illinois Press, 1974, 42-76.
- J.E. Russo. More information is better. Journal of Consumer Research, 1974, 1, 68-72.
- J.E. Russo, G. Krieser & S. Miyashita. An effective display of unit price information. Journal of Marketing 1975, 39, 11-19
- D.L. Scammon. Information load and consumers. Journal of Consumer Research, 1977, 4, 148-155.
- J.O. Summers. Less information is better. Journal of Marketing Research, 1974, 11, 467-468.
- E.F. Taylor. Food processors and open data labeling. National Food Situation, 1976, 155, 27-31.
- J.W. Tukey. Bias and confidence in not quite large samples. Annals of Mathematical Statistics, 1958, 29, 614.
- United States, FTC. Food advertising: Proposed trade regulation rule and staff statement. Federal Register, 39, No. 218, 29842-39862, 11 November, 1974.
- R.H. Warland & R.O. Hermann. The new wave of imitation foods: Problems ahead. Journal of Consumer Affairs, 1971 5, 56-59.
- W.L. Wilkie. Analysis of effects of information load. Journal of Marketing Research, 1974, 11, 462-466.

ALLISON & UHL REVISITED:  
THE EFFECTS OF TASTE AND BRAND NAME ON PERCEPTIONS AND PREFERENCES<sup>1</sup>

G. A. Mauser, Simon Fraser University

Abstract

A 2<sup>3</sup> split-plot factorial experiment was conducted to assess the effect of brand information, product familiarity, and order of presentation on consumers' judgements of the taste of beer samples. Paired comparisons on several taste characteristics and preferences served as dependent variables. Subjects were 240 college students. The results suggest that beer drinkers can distinguish among brands using taste and aroma cues alone.

Introduction

It is frequently asserted that beer drinkers are unable to discriminate among brands of beer using only taste and aroma cues. If true this would imply that beer drinkers are forced to rely on extrinsic cues (e.g., brand name, price, store image) in order to decide which brand of beer to purchase because they can not make use of intrinsic cues (e.g., product taste or aroma). However the empirical support for this claim is somewhat mixed. Only three published studies can be found which have directly addressed this question (Allison & Uhl, 1964; Jacoby, Olson, & Haddock, 1971; and Valenzi & Eldridge, 1973). Two of these three studies support this assertion while the third found that consumers could discriminate among brands using only taste and aroma cues significantly better than chance (Jacoby, et al., 1971).

The failure of two of the studies to find significant differences may have been due to problems in their experimental methods rather than to consumers' inability to discriminate. Close examination of these studies supports this conjecture. Allison & Uhl's pioneering work was a field study which used crude three-point rating scales. Valenzi & Eldridge, while conducting a well-controlled laboratory study, used rating scales which were not subjectively anchored, and they hypothesized that significant differences would have been found had they used better anchored scales.

It is worth noting that these studies differ in their choice of dependent variables. While Allison & Uhl measured a variety of taste characteristics (e.g., strength, body, bitterness), later researchers have focused exclusively on perceived quality. Current work has been primarily on the effect of price on perceived quality (Olson, 1977). There are a few advantages to studying descriptive taste characteristics rather than perceived quality. First, descriptive characteristics allow more scope for marketing control -- they are more "actionable" -- such characteristics are particularly useful in positioning new products (Stefflre, 1968; Green, 1975). Second, it seems plausible that

differences between brands are more readily identified using descriptive characteristics than with perceived quality due to the more "available" or more "codable" nature of descriptive characteristics (Brown & Lenneberg, 1954; Lanz & Stefflre, 1964)

It thus appears worthwhile to replicate Allison & Uhl's earlier study using better experimental controls and more sensitive measurement techniques. The method of paired-comparison is widely regarded as a particularly sensitive measure (Byer & Abrams, 1953; Jellinek, 1964; Torgerson, 1958). In comparison with single stimulus methods (i.e., the rating scale), paired comparison methods two advantages: subjects compare stimuli directly with each other, rather than only indirectly via scale position, and also the task itself makes fewer demands on the subject.

Despite these advantages paired-comparison is rarely used in academic marketing research studies. Perhaps this is due to the large number of pairs of stimuli required with even a modest number of stimuli as well as the forbidding statistical methods required by its use (Scheffe, 1952; Neter & Wasserman, 1974; Bechtel, 1967).

A secondary concern in this paper is to look at the effect of consumer familiarity with the product class on taste judgements. Several researchers have suggested that consumers' perceptions of a product class may differ with their degree of familiarity that they have with the product (Shapiro, 1968; Olson, 1977). Perhaps people who drink beer only occasionally, and thus are less familiar with the product class, are less able to discriminate among the various brands using only taste and aroma cues.

The objectives of this study are to (a) examine to what extent the perception of and preference for beer samples is influenced by brand information, (b) determine which taste characteristics (if any) beer drinkers can use to distinguish among brands of beer using only taste and aroma cues, (c) determine to what extent familiarity with the product class influences drinkers' ability to distinguish among brands of beer.

Method

Stimuli

Nine different brands of beer were included in the study. Brands were selected to cover as broad a range as possible of brand image and composition differences. There were six Canadian brands (two each of ale, lager, and light lager) as well as three American brands (two lager and one light lager).

All of the Canadian brands were exactly the same price while the American brands retailed for a slightly higher price in B.C. All of the brands, Canadian and American, were approximately the same color.

Subjects

A convenience sample of 240 beer drinkers was drawn from volunteer undergraduate and graduate students. Students

<sup>1</sup>An earlier version was presented to the 1977 meeting of the Canadian Association for Administrative Studies in Fredericton, N.B. The author would like to acknowledge the support and funding of Labatt Breweries of B.C., Ltd. and the B.C. Provincial Ministry of Labour. I would also like to thank Jim Watts and Ray Koopman for their help in this study as well as the hard-working students in Commerce 444.

had to be at least 19 years of age and to admit to drinking at least "some" beer to qualify for inclusion in the study. Twenty-seven percent of the subjects in each experimental condition were female. College students were selected for this study because they are an important submarket for the beer industry as well as being a standard group for experimental researchers to study.

#### Design

A 2<sup>3</sup> factorial experiment with one repeated measures factor was conducted having 60 Ss per cell. The repeated measure factor was Label (brand name present or absent). The between-Subject factors consisted of Familiarity (high, low) and Order of Presentation (straight, inverse). Students were classified for the Familiarity factor as having a high degree of familiarity with the domain if they drank 6 or more bottles of beer per week, or as having low familiarity with the domain, if they drank 5 or fewer bottles per week. This break point was determined by asking both Labatt's and Molson's what break point they usually used. The Order of Presentation factor involved reversing for half of the sample both the sequence of pairs that are presented to the S (e.g., p1, p2, p3 or p3, p2, p1) as well as the order of samples within pairs (e.g., in comparing Olympia and Schlitz, Schlitz may be presented either on the left or on the right).

The dependent variables were (a) the perceived similarity of the samples, (b) preference for the samples, (c) judgements of which sample is X'er on each of several taste characteristics (i.e., strength, lightness, after-taste, bitterness, carbonation, heavy bodiedness, fillingness, and smoothness). Judgements of similarity were expressed using a 9-point scale, where 1 indicated "dissimilar" and 9 indicated "very similar". All of the remaining dependent variables were dichotomous choices indicating which stimulus was judged X'er on the characteristic involved. No "undecided" responses were permitted.

Fifteen blocks of three distinct pairs each were fashioned to include all 36 possible pairs of the 9 brands. The nine identity pairs were included in the unlabeled condition, and nine redundant pairs included in the labeled condition, in order to fill out the 9 blocks. There were four replications of each of the 15 blocks in each of the 8 experimental cells.

#### Procedure

Each S tasted three pairs of beer samples under each level of Label (brand name present or absent) yielding a total of six pairs tasted. All Ss received three unlabeled beer samples first and then three labeled beer samples. Each sample of beer contained 1 ounce of beer. Following the tasting of each pair of samples, Ss were requested to indicate their impressions on each of the dependent measures. Ss were requested to nibble unsalted crackers after tasting each pair of beer samples in order to cleanse the palate.

After tasting a pair of samples, Ss responded to each dependent variable in turn. Ss first rated the perceived similarity of the stimuli and then indicated which of the two stimuli they preferred the more. Next Ss judged the samples on each of the eight taste characteristics. These characteristics were presented to Ss in 16 distinct orderings in order to minimize a possible fatigue effect influencing the results. These 16 orderings consisted of all one-step permutations of a pre-selected order and its inverse. Each S was presented with a different ordering of the eight taste characteristics for each of the six pairs of beers he was asked to judge, receiving three permutations from each of the two basic orders in alternation so that a pattern was not

readily apparent in the presented orderings.

Careful attention was paid to presenting all beer samples under identical conditions: the temperature was kept at 39 degrees F (4 degrees C). Bottles and cans were opened and poured immediately before being offered to Ss; the beer samples were carefully poured in a manner that would not create a frothy head.

#### Results

A split-plot ANOVA design was used to analyze these data because of the repeated measure factor (Label). Two 2-way ANOVAs were conducted to examine the effect on the dependent variables by the independent variables of Label, Familiarity, and Order. Two 2-way ANOVAs were conducted rather than one 3-way ANOVA in order to have a sufficient number of Ss (8) per cell for each pair of beer samples (Kirk, 1968, pp.245-318).

In order to examine the effect of the independent variables for the entire set of 9 beer brands, the ANOVAs for each of the 36 pairings were summed for each of the 9 dependent variables. Thus the appropriate degrees of freedom for the overall F-test is the sum of the degrees of freedom for each of the individual F-tests.

Table 1 shows the results of the first two-way split-plot ANOVA for the Familiarity and Label factors for the Ss' overall preferences for the 9 beer brands. Neither of the main effects nor the interaction effect is significant at the .01 level.<sup>2</sup> This would imply that beer drinkers are able to discriminate as well (or as poorly) in taste tests with or without brand information, and that frequent beer drinkers do not differ appreciably from occasional beer drinkers in their preferences.

Table 2 shows the results of the second two-way ANOVA for Ss' preferences in which the Familiarity factor has been replaced by the Order factor. Again, neither of the two main effects are significant, nor is the interaction effect. This provides support for the previous ANOVA results for the Label Effect, and implies that the order of presentation of the stimuli did not play an important role in determining Ss' preferences.

TABLE 1  
LABEL AND FAMILIARITY EFFECTS FOR PREFERENCE MEASURE

Source	Subjects	df	MS	F
Between subjects	139.504	540		
Familiarity	11.121	36	.309	1.212
Subj. w/groups	128.383	504	.255	
Within subjects	131.000	576		
Label effect	11.621	36	.323	1.531
Label x Familiarity	12.996	36	.361	1.711*
Label x subjects	106.383	504	.211	
Total	270.504	1116		

N.B. An F-value of 1.675 is needed for an effect to be significant at the 0.01 level with (36,504) degrees of freedom.

<sup>2</sup>The .01 significance level was selected for use in this study as it is appropriately conservative considering the large number of tests of significance that had to be calculated. If the series of tests for each effect across the nine measures is considered as an "experiment", this gives a probability of .086 that at least one of these

FIGURE 1

LABELLED AND UNLABELLED TORSICA SOLUTIONS FOR JUDGED SIMILARITY OF BEER SAMPLES.

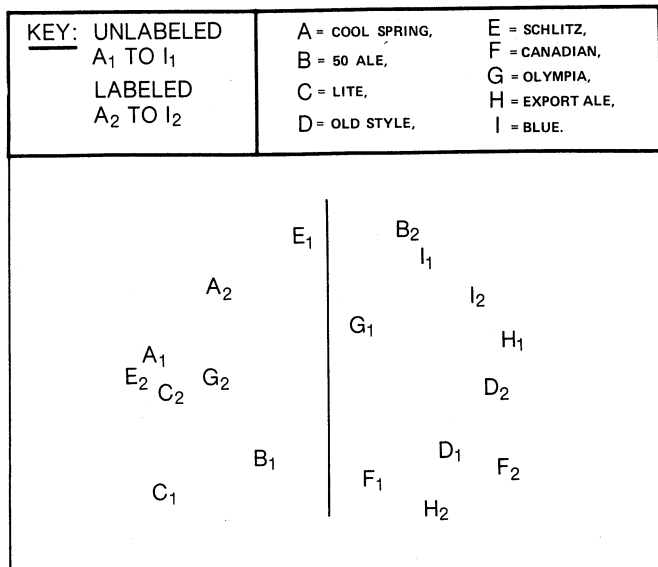


TABLE 2

LABEL AND ORDER EFFECTS FOR PREFERENCE MEASURE

Source	Subjects	df	MS	F
Between subjects	139.504	540		
Order effect	8.371	36	.232	.892
Subj. w/groups	131.133	504	.260	
Within subjects	131.000	576		
Label effect	11.621	36	.323	1.448
Label x Order	6.996	36	.194	.870
Label x subjects	112.383	504	.223	
Total	270.504	1116		

N.B. An F-value of 1.675 is needed for an effect to be significant at the 0.01 level with (36,504) degrees of freedom.

Insignificant ANOVA results may be due to excessive noise in the data or due to the same patterns emerging under both conditions. To investigate the patterns of discrimination among the brands, the perceived similarities of the samples were scales for both conditions of the Label factor (labeled, unlabeled) using TORSICA-9B (Young, 1967). Figure 1 compares the results of these two analyses. The unlabeled two-dimensional configuration (stress = .126) has been rotated to a least squares fit with the labeled configuration (stress = .185) following Pennell & Young (1967). While the beer brands

nine tests would be significant at the .01 level. More striking still is that for the entire series of 27 significance tests, there is a probability of 0.24 that at least one test is significant by chance alone at the .01 level (Kirk, 1968, pp. 77-86).

are tightly clustered in the labeled configuration, they are scattered more in the unlabeled configuration. Nevertheless the same two basic groupings may be identified in both configurations:

- Cluster I : Olympia, Schlitz, Lite, and Cool Spring
- Cluster II : Old Style, Export Ale, Canadian, Blue, and 50 Ale

The vertical line in the figure divides the configurations into these two clusters. Only two brands are mis-sorted (Olympia and 50 Ale), both of which are unlabeled samples. Olympia was seen as being heavier when unlabeled than it was when labeled, and 50 Ale was seen as being lighter unlabeled than labeled. Cluster I contains the lighter and smoother brands, while Cluster II holds the brands seen to be stronger, heavier brands. All of the American beer brands are in Cluster I, although it also contains a Canadian brand (Cool Spring).

In a parallel manner to the analyses of preference, two-way split-plot ANOVAs were conducted for each of the remaining dependent variables, the eight taste characteristics. Tables 3 & 4 show the F-values for these analyses. Neither the Order effect nor the Familiarity effect were significant for any of the taste characteristics. It appears that for the set of measures included in this study, both occasional and frequent beer drinkers agree about the taste of beer. Nor did the order of stimulus presentation play an important role for any of these dependent variables.

TABLE 3

SUMMARY OF F-VALUES FOR TWO-WAY ANOVA'S ACROSS ALL MEASURES FOR FAMILIARITY AND LABEL EFFECTS

Measure	Familiarity	Label	F x L
Preference	1.21	1.53	1.71 *
Strength	0.59	2.17 *	0.94
Lightness	1.20	2.49 *	1.05
After taste	1.13	1.15	1.19
Bitterness	1.37	1.29	1.20
Carbonation	1.02	0.99	0.93
Heavy Bodiedness	0.58	2.35 *	1.18
Fillingness	1.02	1.80 *	0.92
Smoothness	1.07	1.04	1.19

N.B. An F-value of 1.675 is needed for an effect to be significant at the 0.01 level with (36,504) degrees of freedom.

There was a significant Label effect in four out of the eight taste characteristics. The introduction of brand information appears to have altered consumer perceptions of beer strength (F = 2.13, p < .01), lightness (F = 2.49, p < .01), heavy bodiedness (F = 2.28, p < .01), and fillingness (F = 1.78, p < .01). Interestingly enough, brand information did not seem to influence perception of aftertaste, bitterness, carbonation, nor smoothness for these beer brands.

Only one of the interaction effects was found to be significant, that of Familiarity x Label (F = 1.71, p < .01). This interaction was also found to be significant in Valenzi and Eldridge (1973).

TABLE 4

SUMMARY OF F-VALUES FOR TWO-WAY ANOVA'S ACROSS ALL MEASURES FOR ORDER AND LABEL EFFECTS

Measure	Order	Label	O x L
Preference	0.89	1.45	0.87
Strength	0.80	2.13*	0.71
Lightness	1.30	2.60*	1.53
After taste	1.08	1.13	0.90
Bitterness	0.91	1.27	0.96
Carbonation	1.29	1.02	1.24
Heavy Bodiedness	0.82	2.28*	0.77
Fillingness	1.22	1.78*	0.83
Smoothness	0.95	1.03	0.95

N.B. An F-value of 1.675 is needed for an effect to be significant at the 0.01 level with (36,504) degrees of freedom.

How robust are these results considering that the paired comparisons are not strictly independent with respect to each other as every brand of beer is compared with every other one? To attempt to answer this query, the ANOVA's were recalculated omitting, one at a time, each of the nine brands, leaving 28 paired comparisons for the remaining eight brands. Not one of the previously significant effects was now insignificant, nor were any of the previously insignificant effects now significant. This suggests that these results are not an artifact of the lack of independence.

It is also possible that the two-way ANOVA's used in this study did not find significant effects since the sums of squares had been collapsed too far by including variance attributable to the effect of the third factor in the error terms. To evaluate the importance of this potential problem, the preference measure, which was not significant but was quite close, was reanalyzed using a 3-way ANOVA design. None of the main effects were found to be significant at the .01 level, although the Familiarity x Label interaction effect was found to be significant as it was in the 2-way analysis ( $F = 1.77$ ,  $p < .01$ ). None of the other measures would be expected to yield different results if reanalyzed using a 3-way ANOVA.

TABLE 5

THREE WAY ANALYSIS OF VARIANCE FOR PREFERENCE MEASURE

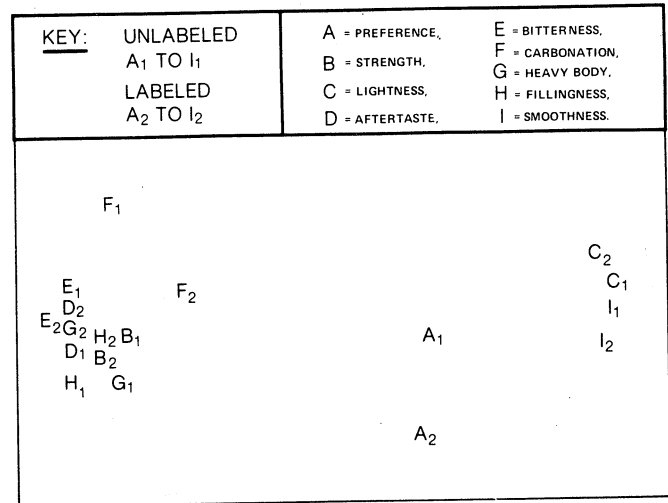
Source	Subjects	df	MS	F
Between subjects	134.629	540		
Familiarity	11.121	36	.308	1.253
Order	6.746	36	.187	.760
F x O	10.246	36	.284	1.154
Subj. w/groups	106.500	432	.246	
Within subjects	135.000	576		
Label	11.746	36	.326	1.484
F x L	13.996	36	.388	1.768*
O x L	8.871	36	.246	1.120
F x O x L	5.371	36	.149	.678
L x subj. w/groups	95.000	432	.219	
Total	269.629	1116		

N.B. An F-value of 1.678 is required for an effect to be significant at the .01 level with the degrees of freedom (36,432).

That the label factor was significant in four out of the nine measures suggests that brand information influences certain taste characteristics but not others. This is probably due to the information associated with the brand images which pertains to specific taste characteristics but not to others. These results are consistent with the often heard claim that American beers are lighter and have a lower alcohol content than Canadian beers. However, certain Canadian beers (e.g., Cool Spring and Blue) are judged from taste and aroma to be nearly as light as some American brands (e.g., Schlitz and Olympia).

FIGURE 2

TWO DIMENSIONAL TORSCA SOLUTION FOR INTER-CORRELATED THURSTONE SCALES



To investigate the interrelationships between the measures with respect to the two levels of the Label factor, each of the taste characteristics and overall preference were scaled using Thurstone paired-comparison techniques and the resulting one-dimensional scales inter-correlated (Torgerson, 1958). These intercorrelations were then themselves used as input to TORSCA (Young, 1967). Figure 2 shows the pattern of relationships among the 18 resulting Thurstone scales (9 for each of the two conditions) in the two-dimensional configuration (stress = .056). Note that all scales cluster quite tightly in two distinct groupings except for the two preference scales and, to a lesser extent, the two carbonation scales. These two groupings are due to the very high intercorrelations among almost all scales coupled with the inverted ordering of smoothness and lightness with respect to almost all of the other scales. Note also that both the labeled and unlabeled versions of each scale are located very close to each other with the striking exceptions of the preference and carbonation scales.

## Discussion

In sharp contrast to Allison and Uhl's study, the results of this study suggest that beer drinkers can distinguish among major brands of beer using only taste and aroma cues. Beer drinkers were able to use four of the eight taste characteristics in the unlabeled condition to discriminate among beer samples nearly as well as they could in the labeled condition. Moreover, consumers' preferences did not significantly change from the unlabeled to labeled conditions.

There are a few alternative ways to approach the question of why these studies differ. The first is methodological.

The failure of Allison and Uhl to find significant differences in the unlabeled condition may have been due to the cruder controls that are possible in a field study as opposed to a laboratory study. Moreover, the method of pair comparisons that was used here is more sensitive to existing differences than is the method of single stimuli. The method of single stimuli is particularly insensitive if a very few rating categories are used, as was the case in Allison & Uhl's study.

Other methodological differences which may be able to explain the different results of the two studies should also be considered. First, despite the approximately identical color of the beer samples, some differences may have been observable to the subjects as they were not blindfolded for the tests. Second, the subjects in this study were college students, while Allison & Uhl used adult beer drinkers. Greater confidence may be obtained in the generality of these results if this study were to be replicated using blindfolded adult beer drinkers.

An alternative approach to understanding why these studies might differ is to examine the range of the stimuli used in the two studies. The range of taste differences was larger in this study than was the case in Allison & Uhl's study. Clearly Ss would be expected to be able to distinguish between tastes that were very different using only taste and aroma cues, say between pure alcohol and pure water, or between Guinness and Beaujolais. The interesting question is which differences are beer drinkers able to identify, not whether or not they can identify any differences. All of the brands in earlier studies were American brands, while in this study both Canadian and American brands were included. Moreover, this study found only two clusterings among the brands and all of the American brands fell into the same cluster.

#### Appendix

Fifteen blocks of pairs of beer brands are shown here for both levels of the Label factor. Each S was presented with the task of drinking the beer samples in an assigned row.

Brands Not Identified	Brands Identified
1. (a,b) (b,g) (e,e)	(a,b) (b,g) (b,i)
2. (a,c) (b,d) (b,b)	(a,c) (b,d) (d,i)
3. (a,d) (b,e) (d,d)	(a,d) (b,e) (a,i)
4. (a,e) (d,f) (a,a)	(a,e) (d,f) (c,i)
5. (a,f) (d,g) (g,g)	(a,f) (d,g) (b,g)
6. (a,g) (d,h) (g,g)	(a,g) (d,h) (c,g)
7. (c,h) (d,i) (h,h)	(c,h) (d,i) (a,g)
8. (a,i) (c,d) (c,c)	(a,i) (c,d) (d,g)
9. (c,e) (d,e) (i,i)	(c,e) (d,e) (b,h)
10. (c,f) (a,h) (f,h)	(c,f) (a,h) (f,h)
11. (c,g) (b,c) (f,g)	(c,g) (b,c) (f,g)
12. (b,i) (b,h) (g,h)	(b,i) (b,h) (g,h)
13. (c,i) (b,f) (f,i)	(c,i) (b,f) (f,i)
14. (e,f) (e,i) (g,i)	(e,f) (e,i) (g,i)
15. (e,h) (e,g) (h,i)	(e,h) (e,g) (h,i)

#### References

Ralph I. Allison and Kenneth P. Uhl, "Influence of Beer Brand Identification on Taste Perception", Journal of Marketing Research, 1 (August 1964), 36-39.

G. G. Bechtel, "The Analysis of Variance and Pairwise Scaling", Psychometrika, 32 (1967), 47-66.

Roger Brown and Eric Lenneberg, "A Study in Language and Cognition", Journal of Abnormal and Social Psychology, 49 (1954), 454-462.

Albert J. Byer and Dorothy Abrams, "A Comparison of the Triangular and Two-Sample Taste-Test Methods", Food Technology, 7 (1953), 185-187.

Paul Green, "Marketing Applications of MDS: Assessment and Outlook", Journal of Marketing, 39 (January 1975), 24-31.

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), 570-579.

Gisela Jellinek, "Introduction to and Critical Review of Modern Methods of Sensory Analysis", Indian Journal of Nutrition and Dietetics, 1 (1964), 219-260.

Roger E. Kirk, "Experimental Design Procedures in the Behavioral Sciences", (Belmont, CA: Wadsworth Publishing, 1968).

De Lee Lanz and Volney Stefflre, "Language and Cognition Revisited", Journal of Abnormal and Social Psychology, 69 (November 1964), 472-481.

John Neter and William Wasserman, Applied Linear Statistical Models, (Homewood, Ill.: Irwin, 1974).

Jerry C. Olson, "Price as an Informational Cue: Effects on Product Evaluations", in Consumer and Industrial Buying Behavior, Arch Woodside, Jagdeth Sheth and Peter Bennett (Eds), (Amsterdam: North Holland, 1977), 267-286.

Roger J. Pennell and Forrest W. Young, "An IBM system/360 Program for Orthogonal Least-Squares Matrix Fitting", Behavioral Science, 12 (March 1967), 165.

B. P. Shapiro, "The Psychology of Pricing", Harvard Business Review, 46 (December 1968), 4-16.

Henry Scheffe, "An Analysis of Variance for Paired Comparisons", Journal of the American Statistical Association, 47 (1952), 381-400.

Volney J. Stefflre, "Market Structure Studies: New Products for Old Markets and New Markets (Foreign) for Old Products", in Frank M. Bass, C. W. King and E. A. Pessemier (Eds.), Application of the Sciences in Marketing, (New York: Wiley, 1968), 251-268.

Warren Torgerson, Theory and Methods of Scaling, (New York: Wiley, 1958).

Enzo Valenzi and Larry Eldridge, "Effect of Price Information, Composition Differences, Expertise, and Rating Scales on Product-Quality Rating", Proceedings of the 81st Annual Convention of the American Psychological Association, 8 (1973) 829-830.

Forrest W. Young and Warren Torgerson, "TORSCA, a FORTRAN IV Program for Shepard-Kruskal Multidimensional Scaling Analysis", Behavioral Science, 12 (November 1967) 498.

PUPIL DILATION MEASURES IN CONSUMER RESEARCH:  
APPLICATIONS AND LIMITATIONS

David C. Arch (Student), University of California, Los Angeles

Abstract

Potential applications of pupil dilation measures to consumer research are reviewed. Although the method is a valuable measure of mental activity in consumer tasks, pupil dilations have limitations that will preclude their widespread use.

Introduction

Eight years have elapsed since the last published review on pupil dilation measures in consumer behavior (Blackwell, Hensel, and Sternthal, 1970). The authors concluded that pupil dilations, temporary changes in pupil size, may be an effective measure of consumer response to advertising, but were unsure what processes the dilations actually measured. More recently, several trends have been noted:

- a call by a number of researchers (Payne and Ragsdale, 1978; Russo, 1978) for joint methodologies to study consumer information processing leading to product choices. The objective of multiple techniques is to improve the accuracy of process measurement.
- increased attention to eye fixations as a process-tracing device in consumer behavior (Russo and Rosen, 1975).
- a more definitive interpretation of what pupil dilations actually measure (Goldwater, 1972).

Based on these trends, it is important to update the findings of Blackwell, et. al., to determine if pupil dilations are appropriate and practical measures for consumer behavior experimentation. While pupil dilation measures might fill a void in the capabilities of current process-tracing methods, limitation of pupil dilation measures will likely prevent their widespread use.

The Need for Additional Process-Tracing Methods

A recent review by Russo (1978) discusses five process-tracing methods currently being used in consumer behavior: chronometric analysis, eye fixations, information boards, input-output analysis, and verbal protocols. One apparent weakness in all the current methods is the failure to measure the importance consumers place on individual information inputs (ads, information board cards, etc.) as a basis for their response or choice. In fact, most methods rely on indirect measures of importance, such as the subject's self-report, or avoid the question of importance entirely. Without a true measure of importance, researchers are at a loss to explain the impact of information, one of the areas where consumer behavior studies should have important inputs for marketers and policy-makers.

Pupil Dilation

Eye pupils have been shown to dilate in response to autonomic activity in the nervous system, with dilations ranging from 1.5 to over 9.0 millimeters. While pupil dilations have been used in a number of psychophysical experiments (see Goldwater, 1972, for a review), applications in consumer behavior have been limited and largely unsuccessful.

Edgar Hess and others (Hess, 1965; Krugman, 1965) have employed pupil dilations as a measure of consumer affect to advertising. The pupils were thought to dilate in response to pleasant stimuli and contract when viewing unpleasant stimuli. Subsequent studies (Woodmansee, 1965) have found no direct relationship between pupil dilations and affect.

King (1972) used pupil size and eye direction of models depicted in advertisements as an independent variable in ad copy effectiveness study. While King's manipulations were significant, the experiment does not provide insights on the use of subject pupil dilation as a dependent variable in consumer behavior.

Meanwhile, psychophysiology experiments were establishing pupil dilation as a measure of the subjects' mental effort expended in cognitive tasks. Hess and Polt (1964) found that peak dilations of subjects was related to the difficulty of arithmetic tasks.

Kahneman and Beatty (1966) confirmed the findings of Hess and Polt in a different application, short-term memory tasks. Subjects' pupil size was sensitive to the amount of information being processed. The researchers proposed that variations of pupil size resulted from rehearsing a series of digits or words that would later have to be recalled.

Beatty and Kahneman (1966) applied similar pupil size measurements in tasks involving "more permanent" memory. Processing stages were observed in the experimental task, with peak dilations immediately following the presentation of the experimental task, and a return to normal pupil size after the stimulus was reported by the subject. Due to the wider range of dilations observed, the researchers concluded that longer-term memory tasks required greater mental effort than activities utilizing short-term memory. The subject's range of dilations was shown to diminish following extensive repetitions of a simple task, showing that learning can reduce cognitive demands.

The above studies and others cited in Goldwater (1972, p. 345) have provided support for the notion that pupil dilations reflect mental effort or activity. If mental activity could be equated with importance of the stimulus to the individual, pupil dilations would be a valuable addition to existing process-tracing methods in consumer behavior. But there are several apparent drawbacks:

- Although mental activity has important associations with importance, the relationship is not complete. For example, an unimportant stimulus could evoke substantial mental activity if it was difficult to perceive and understand. The reverse may also be true. An important stimulus might require little mental activity if it was already well-learned by the subject through prior repetition.

• Pupil dilation measures simply reflect mental activity. In themselves, no clue is provided as to what is being processed, but rather that processing is taking place. Subject protocols would also be required if the researcher wants to measure the meaning, as well as the amount, of mental activity.



One application of pupil dilation to consumer behavior does not require the effort-importance link. Public policy is concerned with optimum product information presentation to consumers. Pupil dilation measures could indicate the relative difficulty consumers experience in understanding different information presentation formats.

#### Design Problems in Experiments Employing Pupil Dilation Measures

Like all other physiological measures, pupil dilations are subject to a variety of distortions and interpretation problems. Tryon (1975) discusses twenty-three sources of pupil size variation, with no indication that the listing is complete. The following considerations are of particular importance:

- The light reflex is the strongest determinant of pupil size. Significant variations in the lighting used for the experiment would render the dilation results meaningless. As a pre-measure, the subject's pupil dilation should be measured under standard light changes.
- Subject factors such as age, pupil unrest (hippus), drowsiness, alertness, arousal, and physical measures such as heart rate and blood pressure have demonstrated effects on dilation. The experimenter must consider these factors in subject selection, and monitor physical signs during the experiment to aid in the interpretation of results.
- It is suggested that dilation experiments be conducted at roughly the same time of day, using subjects with the same number of hours sleep. The experimental sessions should be as brief as possible to control for subject habituation to the task.
- The amount of incentive, the degree of motor or verbal response required, and the information load resulting from the instructions to the subject should be held constant.
- The lid closure (blinking) reflex results in a momentary pupil contraction followed by a dilation. The subject should be told when to blink to control for the distortion of results.

Woodmansee (1966) states that a delayed dilation response to a stimulus may change the baseline for measuring the pupillary effects of the next stimulus. Clearly, a separation between stimulus presentation must be maintained if the response to each stimulus is to be measured.

Simpson and Paivio (1968) have found that dilation effects are increased if motor responses or verbal protocols are required concurrent with the stimulus presentation. Unless the experimenter wishes to measure the total dilation for all activities or has a proven method of partialling out these effects, any motor or verbal activity should be separated from stimulus presentation.

In summary, effective use of pupil dilation measures in a consumer behavior experiment involving visual stimuli should include:

- A constant light source. Videotape monitors or a cathode ray tube (CRT) device would be effective for this purpose, since both permit contrast adjustments to a mid-range of illumination.
- Consistent presentation formats for visual stimuli. Typed pages or transparencies would be acceptable, if the same order and organization is used for all stimuli.
- Separation of presentation and response, so that the

individual dilation effects can be measured. Intervals allowing for return of pupils to baseline levels would aid in the interpretation of results.

- Control of subject factors prior to and during the experiment. Careful selection of a reasonably homogenous subject group, and proper instructions as to desired eye fixations and blinking are essential.

A high degree of equipment sensitivity is required for proper monitoring of pupil dilations. Beatty and Kahneman (1966) observed average pupil size changes of 1.1 millimeters or less in their cognitive load studies. Pupil magnification would help to correct, but not eliminate, the problem of minute measurement.

Proper monitoring of pupil dilation also requires measurement. Several types of pupillometers, commercial devices that measure pupil size, are suitable for pupil dilation experiments. Pupillometers are available that record pupil size from 20 to 30 times per second, a measurement frequency used in current physiological experiments. One manufacturer produces a device that simultaneously measures eye fixation and pupil dilations.

A dedicated computer is required to consolidate the idea produced by the dilation experiment. Several technical computers can be used in conjunction with a pupillometer, and are capable of presenting experimental output in a variety of forms ranging from raw data on pupil size to computer-plotted graphs.

#### Potential Applications of Pupil Dilation in Consumer Research

Several applications of pupil dilation in consumer behavior experiments are proposed as examples of the potential utility of the measure:

- As a measure of mental activity associated with subject's cognitive responses to persuasive messages. (Wright, 1973).
- As a processing effort measure associated with verbal protocols. Subject reports of their processing activity can be associated with pupil dilations on an "on-line" basis (Payne and Ragsdale, 1978) or in the form of a "prompted protocol," where the subject verbalizes his/her thoughts relative to a previous task (Russo and Rosen, 1975).
- As a mental activity measure associated with successive fixations in eye movement analysis (Russo, 1978).

These applications are tentative at best, with a host of experimental design questions yet to be solved.

#### Discussion

Although pupil dilation measures have potential value in consumer research, limitations of the method are also apparent. Measurement of pupil size is an exacting science, and few researchers can afford the equipment required. Even with the most sophisticated devices, experimental conditions would have to be exceedingly precise to avoid confounding the dilation results. Verbal stimuli are preferred to visual presentations in order to avoid light reflex effects.

Obtrusive measurement of pupil size is not considered to be a problem, since pupil dilations are veridical representations of mental activity. But pupil dilation measures are totally impractical in any setting but the laboratory, restricting their research applications to lab experiments.

Most researchers, therefore, will probably settle for "second-best" measures of importance and mental activity, or beg the question entirely. Still, the potential benefits of pupil dilation measures might warrant further feasibility studies in consumer research.

J.J. Woodmansee (1965), An evaluation of pupil response as a measure of attitude toward Negroes, Doctoral Dissertation, University of Colorado.

Peter L. Wright (1973), "The Cognitive Processes Mediating Acceptance of Advertising," Journal of Marketing Research, X, 53-62.

#### REFERENCES

Jackson Beatty and Daniel Kahneman, (1966), "Pupillary changes in two memory tasks," Psychonomic Science, 5, 371-2.

Roger D. Blackwell, James S. Hensel and Brian Sternthal, (1970), "Pupil Dilation: What Does It Measure?" Journal of Advertising Research, 10, 15-18.

Roger D. Blackwell, J.S. Hensel, M.B. Phillips and B. Sternthal (1970), Laboratory Equipment for Marketing Research, Dubuque, Iowa: Kendall Hunt Publishing Co.

James F. Engel, R.D. Blackwell and David T. Kollat (1978), Consumer Behavior, Hinsdale, Ill.: Dryden Press, 352-53.

Bram C. Goldwater (1972), "Psychological Significance of Pupillary Movements," Psychological Bulletin, 77, 340-55.

E.H. Hess, and J.M. Polt (1964), "Pupil size in relation to mental activity during simple problem solving," Science, 143, 1190-92.

E.H. Hess (1965), "Attitude and Pupil Size," Scientific American, 212, 46-54.

Daniel Kahneman and Jackson Beatty (1966), "Pupil Diameter and Load on Memory," Science, 154, 1583-85.

Albert S. King (1972), "Pupil Size, Eye Direction, and Message Appeal: Some Preliminary Findings," Journal of Marketing, 36, 55-58.

H.E. Krugman (1965), "A comparison of physical and verbal responses to television commercials," Public Opinion Quarterly, 29, 323-24.

John W. Payne, E.K. Easton Ragsdale (1978), "Verbal Protocols and Direct Observations of Supermarket Shopping Behavior: Some Findings and a Discussion of Methods," in Advances in Consumer Research, Vol. 5, ed. H. Keith Hunt, Association for Consumer Research.

Berkeley Rice (1974), "Rattlesnakes, French Fries, and Pupil-ometric Oversell," Psychology Today, 55-59.

Edward J. Russo (1978), "Eye Fixations Can Save the World; A Critical Evaluation and a Comparison Between Eye Fixations and other Information Processing Methodologies," in Advances in Consumer Research, Vol. 5, ed. H. Keith Hunt, Association for Consumer Research.

Edward J. Russo and Larry D. Rosen (1975), "An Eye Fixation Analysis of Multi-alternative Choice," Memory and Cognition, 3, 267-76.

Herb M. Simpson and Allen Paivio (1968). "Effects on pupil size and manual and verbal indicators of cognitive task fulfillment," Perception and Psychophysics, 3, 185-90.

Warren Tryon (1975), "Pupillometry: A Survey of Sources of Variation," Psychophysiology, 12, 90-93.

J.J. Woodmansee (1966), "Methodological problems in pupillographic experiments," Proceedings of the 74th Annual Convention of the American Psychological Association, 1, 133-34.

## BACK TO THE BASICS: DISCUSSION

James F. Engel, Wheaton Graduate School (Illinois)

### Introduction

In a sense, the discussant in this session has a difficult assignment. The title, *Back to the Basics*, probably is an accurate one, but it is not altogether clear that this was a deliberate intent of those who presented papers. Rather, it reflects more on the difficulty faced by the program chairman in finding themes to use in organizing a multitude of sessions. Therefore, the common thread between the papers is not especially evident.

Perhaps one can stretch a bit and isolate these "basics" in the papers which were presented: (1) methodological validation (Arch and Kendall and Fenwick) and (2) replication (Mauser). In a limited sense, each has shed some light on consumer behavior research in these respects.

### PDR Measures (Arch)

About 1964, Herb Krugman, then with Marplan, burst on the scene with the latest great breakthrough--pupil dilation research. Glowing reports on this daring innovation soon were published by Krugman (1964 and 1965). The discussant also was with Marplan during that period and had opportunity to delve deeply into this methodology and found little in the way of conclusive validation. Later research has shown pretty clearly that the vast investment required for such equipment does little or nothing to augment present advertising research methodology (Goldwater, 1972 and Rich, 1974). In fact, it can lead to some misleading conclusions. Rice reports the example that designers of a frozen french fries ad were delighted with PDR test results until further analysis showed that people were looking at the steak in the picture rather than french fries (Rice, 1974).

David Arch has given us a competent review on this subject, and his paper is helpful for that reason alone. He has not, however, gone much beyond the contributions of Blackwell, Hensel, and Sternthal (1970). It was known at that time that PDR apparently measures load processing and that it says little about the importance of stimuli to the individuals in question. The pioneering research undertaken at The Ohio State University gave pretty strong indications that GSR (galvanic skin response) measures probably should be used in tandem with PDR on the likelihood that GSR measures arousal and importance. Arch unfortunately has not alluded directly to this research, because it is one way out of the dilemma that he has isolated.

The greatest contribution of this paper is that it clearly spells out the difficulties in using PDR methodology which must, by necessity, be confined to highly artificial laboratory settings. These considerations should be a red light to those who wax too enthusiastically over this type of physiological measure.

This reviewer is most puzzled by the author's conclusion that PDR measures have potential value in consumer research. As justification he cites three examples which are, without exception, confined to academic research. What practical applications exist, if any? It was Krugman's overly-enthusiastic claims on this dimension that ignited a controversy during the 1960s. PDR has been proposed and apparently used as a part of pretesting methodology in advertising. Is this valid or not? Arch must address such a pragmatic consideration; otherwise his paper is not much more than a competent review of what others have already said. It would seem from his review that PDR is pretty much of a dead end in all respects. If this is what he wanted to say, then it should be stated loud and clear in a pragmatic context of practical applications.

### The Shopper and the Label (Kendall and Fenwick)

Kendall and Fenwick have placed a heavy burden on the reader or listener as they wade through what seems to be an excessively difficult-to-decipher manuscript. But the effort is worth it for one significant reason--some really valuable new light is shed on the validity of self-reports in information processing.

These researchers wisely suspected that laboratory or self-report survey data may obscure the real situation with respect to usage of product labels by supermarket shoppers. Therefore, they did the only reasonable thing by designing a field study centering mostly on observation of actual behavior. In a very real sense, they have provided a validity check on the other methods and found them to come up severely wanting. Their data would have been even more conclusive had they compared actual observation and self-reports with the same people studied in their project. This probably is the next step to be taken by those interested in further research.

The other data reported are only of incidental interest. For example, it would have made much more sense to have probed immediately after observation to try and uncover which information was used. The label ranking experiment was artificial and of doubtful validity. Moreover, the attempt to classify shoppers by discriminant analysis made use of correlated variables and hence is questionable. For example, the type of shopper and the time of day of shopping are obviously interrelated.

Putting these considerations aside, Kendall and Fenwick have helped us by their careful attempt to remedy methodological deficiencies. In so doing, it appears that conventional wisdom about label use is defective. It is by such efforts as this that progress is made in arriving at defensible principles and propositions.

Any field within the behavioral sciences progresses toward maturity by developing a tradition of rigorous replication. Blackwell, Kollat, and the discussant have long called for replication in their books and at previous ACR Conferences. It is encouraging to see in very recent years that replication is underway, and this paper should be viewed in that context.

There is no question that Mauser improved on Allison and Uhl's design by making use of paired comparisons. Allison and Uhl, however, did have an advantage of allowing people to sample various brands of beer, both labeled and unlabeled, in a more natural setting. That fact in itself, of course, removed some of the opportunities for control that Mauser instituted. His bottles consumed perweek is arbitrary and certainly is no improvement on Allison and Uhl's criterion (beer consumed at least three times per week). Other than that, the controls on the actual experimental conditions, the analysis of order bias, and the use of perceived similarity measures are genuine improvements.

There is one place in which Mauser appears to be measuring something different from Allison and Uhl who rated beers along various attributes in terms of "too much," "just enough," and "not enough." Mauser, on the other hand, merely focused on whether one of the samples in each pair was judged as having more of the particular characteristic. While he was following standard paired comparison procedures, his data are not really comparable to Allison and Uhl's, thus making direct comparison difficult.

Contrary to Mauser's own statement, it can hardly be said that his findings are in "sharp contrast" to those of Allison and Uhl. Drinkers were able to distinguish between the eight brands on only four of the eight attributes in both experimental conditions. As he indicated, this is quite likely due to the wider range of tastes in the beers used. If that is the case, then how is this a replication? Furthermore, the perceived similarity ratings were only roughly comparable between the two experimental conditions. The author did not do justice to this latter analysis, because one cannot help but wondering whether other factors also were shaping these perceptual judgments. In short, the case has not been proved definitively.

But suppose there had been really major differences between the two experiments? Many will ask, so what? Was the author merely replicating an interesting experiment, or was he testing a proposition of consumer behavior? Nothing is said about the implications in terms of consumer information processing. As it stands, one could argue that a cannon was used to kill a flea. Allison and Uhl were saying something important in terms of a conceptual framework, and that seems to have become lost here. Replication is of value only when these conceptual issues are given primary focus. Because this focus is absent here, at least insofar as the writeup, an otherwise useful paper probably is destined to fall into the "interesting but so what?" category.

Of these three papers, Kendall and Fenwick made the greatest contribution through their clearcut challenge to existing methodology. As was mentioned, their case would have been even stronger had there been direct methodological comparison, but no matter. They have issued a real challenge both to existing methodology and to the conclusions which have appeared in the literature. The door which has been opened should be used by others, keeping in mind that we always should strive toward "real world" behavioral conditions while endeavoring to maintain experimental control. Kendall and Fenwick showed that this is by no means impossible.

That paper by David Arch is most notable for its careful literature review, although he did neglect some important substantive considerations as noted. Nevertheless, he has demonstrated the value of assessing findings from fields beyond the borders of consumer research. Such scholarly cautions could lead us to avoid the promotional overkill which always seems to accompany a tool such as PDR. It is worth mentioning once again that the inflated claims during the 1960s never had any real substance in fact. Yet those without methodological sophistication piled on the bandwagon. Now we find that Krugman is at it again, this time claiming that analysis of brainwaves is the great new frontier again (Krugman, 1978). While he could be correct, we are in danger of making the same mistake made in the 1960s unless some of us react with healthy skepticism and put on the caution light as Arch has done with PDR. The questions we should ask are these: (1) "what has been learned elsewhere?" and (2) "how does technique represent any improvement over existing methodology?" This requires both extensive literature review and validation research.

Finally, replication is a necessity, and Mauser should be commended in this sense. But let's get some sense of priority and avoid replication just for the sake of replication. The emphasis should be on replication to verify or disprove important propositions. This, in turn, demands a grasp of those areas in which this type of verification is most needed.

#### References

- Roger D. Blackwell, James S. Hensel and Brian Sternthal (1970), "Pupil Dilation: What Does It Measure?" Journal of Advertising Research, 10, 15-18.
- Bram C. Goldwater (1972), "Psychological Significance of Pupillary Movements," Psychological Bulletin, 77, 340-55.
- Herbert E. Krugman (1964), "Some Applications of Pupil Measurement," Journal of Marketing Research, 1, 15-18.
- Herbert E. Krugman (1965), "A Comparison of Physical and Verbal Responses to Television Commercials," Public Opinion Quarterly, 29, 323-24.
- Berkeley Rice (1974), "Rattlesnakes, French Fries, and Pupil-ometric Oversell," Psychology Today, 55-59.

## DISCUSSION OF "BACK TO THE BASICS" RESEARCH PAPERS

Joseph W. Newman, University of Arizona

At the outset, I should like to comment that all three papers represent conscientious efforts to do careful, thorough work and the two empirical studies attempted to employ improved methodology. This is important, of course, because such efforts are essential to what we hope is continuing progress in understanding consumer behavior.

### The Kendall-Fenwick Paper

I shall start with the paper by Kendall and Fenwick and deal with it more extensively than with the others because it is more closely related to work I have done in the past. The study is interesting for several reasons. It is one of the few efforts to observe what real consumers do in the market place vs. the laboratory. It also is one of the few attempts to compare findings obtained by observations vs. self-reports. The findings are provocative and help highlight a need for more comprehensive research.

The extent to which the authors really found out much about the prevalence of label reading depends, of course, on the validity of their measure of the dependent variable. They made the key assumption that their clockings of the time shoppers spent examining the product alternatives were a function of the extent of reading, processing, and use of label information. That may not be the case. Viewing might represent no information acquisition at all and even if acquired, information may not be used. Among factors which might make for different viewing times among shoppers are skill in label reading which probably is related to experience and education; whether the shopping is done leisurely or in haste; the decisiveness of the shopper; how easy or difficult it is to comprehend information from the label format; and the extent to which there are environmental distractions such as those provided by young children parents may have in tow while shopping.

While such considerations raise questions about the dependent variable, they do not mean that nothing useful can come out of observing behavior. On the contrary, it is important to observe what is happening in real life. No direct measure of reading, processing or use of information is possible from in-store observations alone. Instead one must settle for what hopefully is a useful proxy and the authors were thoughtful in their consideration and use of the time variable.

Perhaps the study's most important contribution is that it provides evidence that measures of search behavior based on direct observation and measures based on self-reports of the same behavior do not agree. Only a few researchers have done this (Jacoby et al., 1976; Newman and Lockeman, 1975). At least two other studies (Olshavsky, 1973; Newman and Lockeman, 1975) found considerably more apparent information seeking activity by using direct observation than one would expect from reported survey research findings. Yet most of what is accepted today as knowledge of amount of consumer information seeking is based on survey data.

The authors stated that the "grabbers" in their self-reports tended to claim more label reading than did the "lookers." It would have been helpful had more of the self-report findings been presented in the paper.

The authors observed a large number of shoppers but there is the usual question about the extent to which the findings can be generalized. No descriptive information was given of the stores' clientele and the products involved in this study may not be representative of most food products. If the stores' customers were upscale on education and income, the reported observation of considerable label reading would not be

as surprising as the lesser amount of reading indicated by the self-reports.

As I am sure the authors are aware, there is danger in generalizing from their limited data on canned meat/fish vs. that on dehydrated soup as to the amount of label reading for established vs. new products because of differences in characteristics of the products and possible differences in their importance to the consumer.

As for the question of who looks at labels, the six explanatory variables used in the first analyses of the data were very limited in nature and really did not add much to understanding, as the authors recognized. Actually, it does not appear that the six variables discriminated better than could be expected by chance. The authors reported that use of the variables resulted in correctly classifying as either "grabbers" or "lookers" 75 per cent of the shoppers. But the ratio of "grabbers" to "lookers" in the sample was one to three. And the test was made using the original data rather than a hold-out sample. Limitations of the variables also are indicated by the finding that the discriminant function for the first six variables showed some sign changes when applied to the Aisle 2 vs. the Aisle 1 products. The other three variables (satisfaction with current label information, extent of support for mandatory federal label regulations, and degree of menu planning) seemed to be more relevant.

An issue is involved in aggregating data. The procedure the authors used assumes that in their evaluations all subjects were trying to do the same thing in the same way. It assumes use of an additive model without interaction, and common subjective utilities for label information. If these assumptions are not valid, aggregating the data can lead to erroneous results. In the study by Best and McCullough (1977) to which the authors referred, subjects' responses to labels first were viewed individually. This led to the identification of three different groups of respondents in terms of subjective utilities for label formats.

In summary, the Kendall-Fenwick study, along with a few others, suggests that we may know less than we have assumed about the amount and character of label reading or, for that matter, consumer information search and use in general. Different methods have yielded different findings. Kendall and Fenwick suggested that for public policy purposes any studies based on self-reported label reading should be treated with caution. The same warning, however, is appropriate for studies employing direct observation. This should not lead us to abandon advocacy of empirical research as an important input to public policy decision making. But right now we must recognize a need for better measures and more complete information on label reading and other information seeking activities. Better data would seem to require more complete data gathering approaches than have been employed thus far.

For data on label reading to be meaningful in terms of consumer welfare, we need to know what the consumer tried to find out and whether the effort succeeded and, if it did, whether it was with ease or difficulty. In this connection, there is a need to distinguish among consumers on the basis of knowledge of the product and brands and other characteristics that make for differences in interest in and ability to acquire and use information. Kendall and Fenwick suggested that it is not as much the type of person but the situational variables which determine label reading. Their data give little support to that conclusion, however, because of limitations of their explanatory variables. Some of the variables may well be proxies for unidentified consumer characteristics we need to know more

about. It is quite reasonable to believe that both consumer characteristics and situational variables may be important.

The suggestion that different information formats may be appropriate for "grabbers" and "lookers" is not very practical in that both shop for the same merchandise in the same stores. More important, the suggestion implies thinking in terms of a fixed dichotomy rather than recognizing that today's "grabbers" may be yesterday's "lookers" who have learned about product and brand alternatives and how to get what they want in the supermarket with a minimum expenditure of time.

The assumption often has been made that more search, more reading and more deliberation mean better informed consumers and, consequently, more satisfactory purchases. Evidence in support of this assumption, however, has been lacking because few, if any, studies have attempted to cover that much ground.

#### The Mauser Paper

The Mauser paper reflects serious concern with employing appropriate methodology in a taste test and care in controlling for such things as order effect and sample presentation. Even so, I must confess to having some difficulty with the procedure employed and with accepting some of the conclusions.

The study is not exactly "Allison and Uhl Revisited" as its title states. There are several differences in addition to methodology. Allison and Uhl confined their test to beer, using six brands, apparently all lager beers. Mauser included three categories of malt beverage: ale, lager beer and light lager beer. This being the case, one would expect that he would find more evidence of consumer ability to tell differences. Allison and Uhl tested to learn whether subjects would rate the brand they said they drank most often as superior in both labeled and unlabeled test conditions. Mauser did not attempt to do that. Another difference worth noting is that Allison and Uhl left a six-pack containing bottles of three unidentified brands with their subjects for one week, then left another six-pack of bottles with labels for another week before picking up rating forms. Mauser used a one-time test of one-ounce samples. The effect of such differences in exposure is a subject for future research.

It would have been helpful had Mauser described more fully the characteristics of the brands and his rationale for brand selection and presented hypotheses or statements as to what he expected to find. If beer consumers can discriminate on the basis of taste and aroma, one might expect that they would perceive brands in each of the three categories (ale, lager and light lager) as being similar. Yet inspection of his Figure 1 reveals that this did not clearly happen. The two ales, for example, were spaced far apart in both the labeled and unlabeled configurations. And Schlitz appeared removed from Old Style and Blue even though all three presumably are lager beers (although the paper did not clearly identify brands by type).

Inspection of Figure 1 as it appeared in his paper also makes it difficult to accept the conclusion of reasonably tight clusters of brands. Instead, one is left wondering what is going on here. The author bases his conclusion that beer drinkers can distinguish among brands on a space diagram which does not seem to make sense even though the stress statistic of .126 for the unlabeled condition is statistically significant at the .05 level for nine stimuli and two dimensions (Klahr, 1969). The stress statistic of .185 for the labeled condition, however, has a .5 probability of being

obtained by chance.

Observations such as I have made cultivate a distrust of the data and a suspicion of an unidentified source of error.

A test such as this one might be more informative if an effort were made to distinguish subjects who can discriminate reasonably well on the basis of taste from those who can not. This could be done by means of a triangle test. It would be interesting to know the proportion of discriminators. A taste test such as Mauser conducted then could use only discriminators as subjects which should eliminate some of the noise which may be clouding the picture presented in Mauser's Figure 1 and the results should be a better guide for product planning.

It appears from Figure 1 that knowing the brand made a difference in how the brand was "tasted" in four of the eight cases (Schlitz, Olympia, 50 Ale and Export Ale). In this connection, it is interesting to note that Allison and Uhl found that the gains in brand ratings which resulted when brand was identified were not uniform among the brands tested. Neither study attempted an explanation. Examination of advertising budgets and messages might provide some insight here.

The author's conclusion that familiarity with the product class did not significantly affect taste judgments is open to challenge because familiarity was not measured. Subjects were categorized as having either high or low familiarity depending on whether they drank six or more bottles of beer a week. So the measure really was of consumption which does not necessarily mean familiarity with the various brands in the product class. A heavy, brand loyal beer drinker may be less familiar with alternative brands than a more occasional beer drinker who engages in considerable brand switching.

As for the finding of a label effect in four taste characteristics, a factor analysis of all eight characteristics used in the ratings might reveal that only one, or perhaps two, dimensions were represented.

In summary, the Mauser findings leave me puzzled and unclear about what contribution the paper makes.

#### The Arch Paper

The paper by Arch is an excellent review of the whole area of pupil dilation. It cites the relevant literature and goes on to succinctly identify potential benefits and limitations of the method, related design problems, guidelines for use and potential applications.

It is clear that much remains to be learned about what pupil dilation actually measures, the relative importance of the problems associated with its use and the sensitivity of dilation to the possible distorting influences. Validity, of course, is the key question here as it was with observations of viewing time in the Kendall-Fenwick study of label reading. Both papers help highlight a need for more research which employs more than one method at a time in order that more may be learned about validity.

Thorough reviews of other research methods like the one Arch has done on pupil dilation would be welcome contributions to our consumer research literature.

## References

Ralph I. Allison and Kenneth P. Uhl, "Influence of Beer Brand Identification on Taste Perception," Journal of Marketing Research, 1 (August 1964), 36-39.

Roger Best and James McCullough, "Evaluation of Food Labeling Policies through Measurement of Consumer Utility," in Keith Hunt (Ed.), Advances in Consumer Research, 1978, 5, 213-219.

D. Klahr, "A Monte Carlo Investigation of the Statistical Significance of Kruskal's Nonmetric Scaling Procedure," Psychometrika, 34, 319-20.

Jacob Jacoby, Robert W. Chestnut, Karl C. Weigl and William Fisher, "Prepurchase Information Acquisition: Description of a Process Methodology, Research Paradigm and Pilot Investigation," in B.B. Anderson (Ed.), Advances in Consumer Research, 1976, 3, 306-314.

Richard W. Olshavsky, "Customer-Salesman Interaction in Appliance Retailing," Journal of Marketing Research, May 1973, 10, 208-212.

Joseph W. Newman and Bradley D. Lockeman, "Measuring Prepurchase Information Seeking," Journal of Consumer Research, December 1975, 2, 216-222.

QUESTIONING THE CONCEPT OF INVOLVEMENT  
DEFINED PRODUCT CLASSES

John L. Lastovicka, Temple University

Abstract

Homogeneity of consumer acquisition behavior is examined within a set of diverse product classes. The degree to which different levels of involvement are related to levels of acquisition behavior is examined.

Introduction

Low involvement theorists (e.g., Krugman, 1965; Ray et al, 1973; and Robertson, 1976) suggest that much of the complexity inherent in studying consumer behavior is due to the broad assortment of products acquired and used by consumers. Accordingly, low involvement theory has categorized product classes and suggested different types of purchase behavior in each product category. In one category are the so called "low involvement product classes." Descriptions of typical purchase behavior for this product category have been different from those as described by the traditional consumer behavior models.

Frequently purchased, commodity-like goods such as toothpaste or canned peas are examples of what have become referred to as low involvement product classes. Following Lastovicka and Gardner (1978), a low involvement product class is one in which most consumers perceive little linkage to their important values and is a product class where there is little consumer commitment to the brands. Less frequently purchased, and more brand differentiated product classes, such as automobiles or stereo equipment are often given examples of high involvement product classes. More direct linkage to personal values and more commitment to brands have been assumed in these latter product classes.

Please note that an involvement defined product classification system is consumer and not product defined. That is, a particular good's inclusion in the low involvement group is not determined by any objective characteristics of the good itself - as in Aspinwall's (1961) "color theory" product classification system. Rather, homogeneous consumer perceptions and behaviors are what determine a particular good's involvement classification.

Low involvement theorists have used the involvement concept as the key variable in explaining differences in behavior across different product classes. In purchasing brands in low involvement product classes, buyer behavior is assumed to be passive, re-active instead of active problem solving. Kassarian (1978a, 1978b) has suggested that Skinner's (1972) behavioral modification approach may be more than enough to explain brand choice in low involvement product classes. The cognitively based traditional models of brand choice behavior are felt to be most appropriate in the high involvement product classes.

The Research Question

The present study is an attempt to question the notion of an involvement based product classification scheme. Two criteria were devised to test the classification system.

Criterion One: Generalizability

A primary concern is of generalizability. In order for any product classification system to be useful, products in a like classification must share similar consumer be-

havior. Products of different classifications should demonstrate heterogeneity. The degree of homogeneity within a given involvement classification (e.g., the homogeneity of behavior with a product in the low involvement classification) is questionable as the involvement product scheme is consumer defined. Given knowledge of the differences in individual values, perceptions and behaviors it is likely that same product behavior may not be similar across individuals.

That is, it is possible that products may not be the object of very similar behavior from individual to individual. If the same product has different meanings to different consumers, then the power of generalizability has vanished. Only if the same product is the object of approximately the same behavior across consumers, can generalizations be made about the product class.

One way of increasing the power of generalizability of the involvement product classification system is to better specify what is being classified. Following Belk (1975), it is recognized that one way to more clearly specify a stimulus is to consider the object-in-situation stimulus. Accordingly, products-in-consumption situations would seem to be a better point of classification than products alone.

Criterion Two: Identification

In order for involvement classification system to be useful, not only must there be homogeneity of behavior across consumers within products-in-situations, but there must be a basis for believing that involvement is the key explanatory variable for the assumed differences in behavior. Identification of the key variable in the classification system is imperative. If involvement is not a key variable, then the terms high involvement product class and low involvement product class are misnomers.

This identification problem severely weakens the power of the involvement classification system. This is especially true in the case when variables, other than involvement, can be used to explain phenomena that involvement supposedly explains.

For example, Hupfer and Gardner (1971) have measured a great disparity between automobiles and toothpaste on an involvement scale. But is it primarily involvement which is responsible for the greater degree of thought and consideration expected to go into an automobile purchase relative to a toothpaste purchase? An equally plausible explanation would rely upon learning or familiarity. With more time between automobile purchase, more consideration should go into such a choice decision since consumers are more likely to forget much of their prior make and model evaluations since the last choice. Moreover, even if memory decay was minimal, market structure - in terms of the makes and models of automobiles available - may have changed considerably since the last purchase. With the toothpaste purchase, less pre-purchase consideration can be explained by greater reliance on the last purchase experience due to increased recency. For this automobile-toothpaste example, the distinction between a reliance on prior learning and re-learning seems as palatable as a distinction based on involvement. This suggests that at least familiarity, a competing key variable, must be tested against involvement.



Sample

During Spring 1978, a systematic sample (Sudman, 1976) of 334 operating household phone numbers was drawn from the Philadelphia Telephone directory. Using a procedure described by Jolson (1977), the telephone subscribers drawn were phoned and a request was made to mail a questionnaire. Permission to mail the questionnaire was not obtained from 76 individuals. From the 258 remaining, 143 useable completed questionnaires were returned. This 143 represents a 43% response rate from the initial sample of 334 and a 55% response rate from those mailed the questionnaire.

Measurement

Low involvement theorists have drawn a sharp contrast in their descriptions of buyer behavior in low involvement and high involvement product classes. For high involvement goods the buyer is characterized as an active problem solver and an information seeker. In contrast, for low involvement products the buyer is seen as a creature of habitual routine. In Howard and Sheth's (1969) terms, the high involvement buyer engages in extensive problem solving while the low involvement buyer engages in routinized response behavior.

Behavior for seven different current or possible future purchase decisions was measured using the extensive problem solving -- routinized problem solving dichotomy. The Appendix shows how the two types of behavior were operationalized in the mail questionnaire. Consumer H, in the appendix, was described as engaged in extensive problem solving while Consumer M was described as a routinized problem solver. Respondents were asked to indicate to what degree they would be like H or M for each of seven purchases. The degree of extensive-routinized response behavior was measured on the following continuum of acquisition behavior:

I am very much like H	I am somewhat like H	Like both H and M	I am somewhat like M	I am very much like M
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

These responses were coded on a 1-5 scale where "very much like H" was set to 1 and "very much like M" set to 5.

The seven product - in - consumption behavior situations were:

- (1) A bottle of wine as a gift for a friend.
- (2) A bottle of wine to be consumed by myself or family at home.
- (3) An automobile for personal use.
- (4) Toothpaste for my personal use at home.
- (5) A loaf of bread to be used by myself or family at home.
- (6) Lightbulbs for use at home.
- (7) A high-fidelity stereo sound system for use at home.

Prior research (Hupfer and Gardner, 1971; Lastovicka and Gardner, 1978; and Belk, 1976) suggest that (1), (3), and (7) are in the high involvement domain while (2), (4), (5) and (6) are in the low involvement domain.

Each respondent's own level of involvement and familiar-

ity for each of the six product classes represented in (1) - (7) was measured with four additional questions. Each of these questions consisted of an item using a 1-5 Disagree-Agree scale. Following Lastovicka and Gardner (1978) involvement is seen as having two components: normative importance and commitment. Normative importance is based on Rokeach's (1975, 1969) value system and the degree to which a product is linked to important values. It was measured with:

"I rate this product as being of the highest importance to me personally."

Commitment is based upon Kiesler's (1971) notion of the bonding or pledging of an individual to a particular choice object. It was measured with this item:

"If my preferred brand or make and model was not available at the store or dealership, it would make little difference to me if I had to choose another brand in this general group of products."

Knowledge about the product class was measured with:

"I think that I could talk about this general group of products for a long time;"

while remembered personal experience with the product class was measured with:

"I can remember having purchased something in this general group of products."

Convergent and divergent validity for the first three items of the four presented above is presented by Lastovicka and Gardner (1978).

Measures were also taken to classify the respondents by sex, marital status, occupation, age and income.

Analysis

Generalizability

Using a repeated measures two factor mixed analysis of variance model (Winer, 1971), the criterion of generalizability was tested. The model examined the influence of individual differences and product-situation combinations on one dependent measure (the response on the 1-5 extensive problem solving -- routinized problem solving scale).

The intent of using analysis of variance (ANOVA) was two-fold. First, it was used to construct estimates of the relative source contributions in the ANOVA model to variance in the dependent measure. Of prime concern was the relative contribution, due to products-situations versus individual differences, on acquisition behavior. Second, ANOVA was used to assess the significance of a contrast between the three product-situations assumed in the high involvement realm and the four product-situations assumed in the low involvement realm.

Using ANOVA in the first manner required stipulating the expected mean squares and the solutions for the estimate of variance components (Winer, 1971; Kirk, 1968). The expected mean squares and the estimates of variance components are shown in Table 1. Since repeated measures within each product-situation treatment were not taken, it is impossible to separate the error component,  $\sigma_e^2$ , from a potential interaction component,  $\sigma_{AB}^2$ . Seeking a simple model, the results presented assume that the interaction component,  $\sigma_{AB}^2$ , is zero.

The results in Table 2 suggest that the products - in -

TABLE 1  
E(MS) AND ESTIMATES OF VARIANCE COMPONENTS FOR ANOVA MODEL

Source of variance	df	E(MS)	Estimate of Variance Component	Percent Contribution
Between People (A)	(n-1)	$\sigma_e^2 + k\sigma_A^2$	$\sigma_A^2 = (MS_A - MS_e) / k$	$\sigma_A^2 / \sigma_A^2 + \sigma_B^2 + \sigma_e^2$
Products -in- Situations (B)	(k-1)	$\sigma_e^2 + n\sigma_B^2$	$\sigma_B^2 = (MS_B - MS_e) / n$	$\sigma_B^2 / \sigma_A^2 + \sigma_B^2 + \sigma_e^2$
Residual (e) (Error + AB)	n-1(k-1)	$\sigma_e^2$	$\sigma_e^2 = MS_e$	$\sigma_e^2 / \sigma_A^2 + \sigma_B^2 + \sigma_e^2$

where: k = number of levels in treatment B  
n = number of subjects

TABLE 2  
RELATIVE CONTRIBUTIONS OF VARIANCE

Source of Variance	df	SS	MS	F	Percent Contribution
Between People (A)	142	556.61	3.92	2.94*	17.37
Products -in- Situations (B)	6	380.24	63.37	47.64*	20.32
Residual (e) (Error + AB)	852	1134.33	1.33		62.35
TOTAL	1000	2071.18			100.00

\*(p < .05)

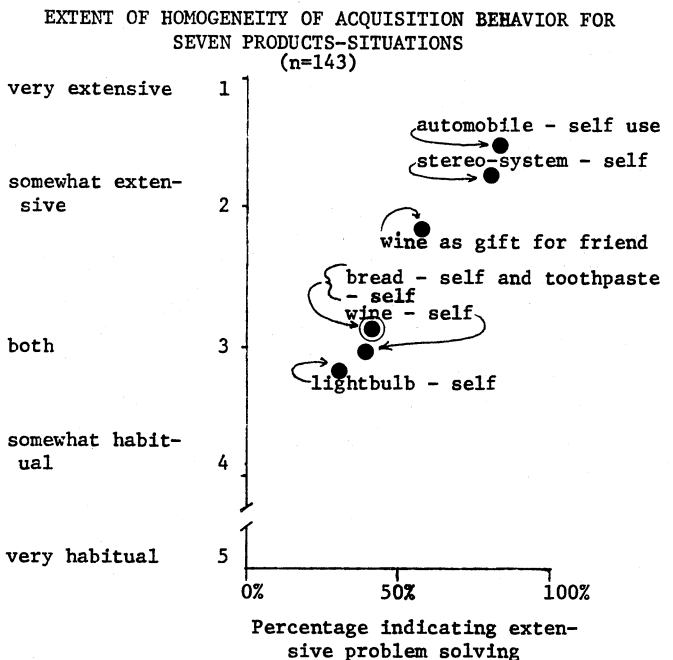
situations do slightly dominate over individual differences in contributing to explained variance. A reasonable interpretation is that differences in acquisition behavior are due primarily to differences in products and how they are to be used and, secondly, to differences between people. Although a large share of the variance, 20.32%, is accounted for by the product-in-consumption-situation notion, a disturbingly large proportion of the variance, 62.35%, is left unaccounted for. This suggests the presence of an interaction component.

Scheffe's multiple comparison test was used to test the difference in mean levels in acquisition behavior between the three so-called high involvement and the low involvement products-in-situations. The mean level for the former group, 2.02, was significantly lower than the mean for the latter group, 3.10, (p < .05). This suggests that average claimed acquisition behavior with the group consisting of automobile for personal use, wine

for gift and stereo for personal use can be best described as "somewhat like" extensive problem solving behavior. For the group consisting of light bulbs for use at home, bread for use by self or family, toothpaste for personal use, and wine for family or self consumption, a description of "like both" extensive and routinized acquisition behavior is appropriate.

Figure 1 is a visual aid in analyzing homogeneity. Figure 1 presents, for each purchase decision, the percentage of the 143 subjects who checked either the 4 or 5 point on the scale (indicating extensive problem solving) versus the mean value of each purchase decision on the same scale. This reinforces the notion that not only are there overall differences in acquisition behavior between purchase decisions, but that the behavior for each type of decision is relatively homogeneous.

FIGURE 1



Identification

So far, seven different products - in - consumption situations have been demonstrated to exhibit different types of acquisition behavior. Furthermore, within products this acquisition behavior is relatively homogeneous across people. The degree to which involvement is the key explanatory variable for these behaviors, must still be tested.

The influence of the two components of involvement, the measures of knowledge and prior experience, as well as a set of demographic measures were used in a multiple regression model to explain the dependent variable of the 5 point scale measure of extensive-routinized response behavior. Since the ANOVA suggests that the respondents were relatively homogeneous in behavior within a product class and heterogeneous across products, there was some justification to consider the data base as consisting of 858 (6 products-personal-use situations X 143 respondents) independent data cases. That is each respondent supplied 6 data cases, one for each product - personal use purchase, into the data matrix used in the multiple regression analysis. The multiple regression model used was:

$$X_1 = B_2X_2 + B_3X_3 + B_4X_4 + B_5X_5 + B_6X_6 + B_7X_7 + B_8X_8 + B_9X_9 + B_{10}X_{10} + B_{11}X_{11} + B_{12}X_{12} + B_{13}X_{13} + B_{14}X_{14} + B_{15}X_{15} + B_{16}X_{16} + B_{17}X_{17} + e \quad (1)$$

$X_1$  = degree of extensive-routinized problem solving claimed for a given product for personal use (the higher the value on the 1-5 scale, the more routinized).

$X_2$  = degree of knowledge about the same given product (the higher the value, the more knowledgeable).

$X_3$  = perception of the normative importance of the same given product (the higher the value, the greater the importance).

$X_4$  = purchase experience with the same given product (the higher the value, the greater the experience).

$X_5$  = non-commitment to brands in the same given product class (the higher the value, the less committed)

$X_6$  = dummy variable indicating sex (0 = male, 1 = female).

$X_7$  = dummy variable for single or never married marital state.

$X_8$  = dummy variable for married marital state

$X_9$  = dummy variable for divorced or separated marital state.

$X_{10}$  = dummy variable for professional occupation.

$X_{11}$  = dummy variable for white collar occupation.

$X_{12}$  = dummy variable for blue collar occupation.

$X_{13}$  = dummy variable for housewife occupation.

$X_{14}$  = dummy variable for student occupation.

$X_{15}$  = dummy variable for retired.

$X_{16}$  = age in years.

$X_{17}$  = income range

(1 = \$5,000 or less, 2 = 5,000 - 10,000, 3 = 10,000 - 15,000, 4 = 15,000 - 20,000, 5 = 20,000 - 25,000, 6 = 25,000 or more)

As a quick review, the variables on the right side of the equation are expected to be related to the dependent measure as follows:

1. Low Involvement Theory Suggests: a negative relationship between normative importance and routinized behavior; and a positive relationship between non-commitment and routinized behavior.
2. Learning and Familiarity Explanations Suggest: a positive relationship between both knowledge and experience and routinized behavior.
3. Other Explanations of Acquisition Behavior Suggest: a positive relationship between age and routinized-like behavior (Hempel, 1969); and a

negative relationship between occupational status and routinized-like behavior (Newman and Staelin, 1972).

4. No particular relationships: for sex, marital status, or income were expected. Since a good deal of acquisition behavior was due to the between people component in the ANOVA, these additional standard measures of individual differences were included in an exploratory manner.

Multiple regression analysis was performed on equation (1) for the 680 data cases of the 858 which contained no missing data. The results of this regression equation are presented in the top half of Table 3. The signs of the beta weights for this equation are generally as expected, except for the very notable reverse of the sign of the knowledge weight. The familiarity hypotheses argues that as knowledge increases, routinized problem solving behavior should increase. The empirical results suggest rather that as knowledge increases routinized problem solving decreases. Another set of unexpected results were the significance of the beta weights for the sex, income and marital state measures. The sex coefficient sign suggests that women are more likely to engage in routinized problem solving while men are more likely to be more careful shopping and engage in extensive problem solving. The income beta weight suggests more routinized behavior for those with higher income. And finally from the net effects of the weights for the marital dummy variables, one may believe that only the widowed state, the fourth marital classification, had influence on acquisition behavior. The beta weight for the retired suggests more routinized behavior for this group.

The interpretation about the effects of widowhood suggests that the effects of the three marital statuses directly modeled in equation (1) - single, married, divorced or separated - are the same. Using a procedure as shown by Rao and Miller (1971, p. 145), this was examined by testing the null hypothesis:

$$B_7 = B_8 = B_9.$$

Since the obtained F (3.38 with 2 and 666 df) was less than the critical value, the null hypothesis stands. Consequently, the only two relevant marital states were widowed and not widowed. This was modeled with a new dummy variable,  $X_{18}$ , representing widowhood.

The independent variables with significant beta weights in equation (1) were retained and used in a second regression equation of the form:

$$X_1 = B_2X_2 + B_3X_3 + B_4X_4 + B_5X_5 + B_{17}X_{17} + B_{18}X_{18} + e \quad (2)$$

The results of this second regression equation are in the lower half of Table 3.

The net effect of the regression results of equation (2) is in general support of using involvement as a partial identifying key in this behaviorally based product classification system. First, the direction of relationships between the two components of involvement, normative importance and commitment, and degree of routinized problem solving behavior are exactly as expected. The expected direction of relationship was not, however, obtained with the measure of knowledge. This weakens the power of this competing explanation. Second, the relative importance of the involvement components as defined by their contribution to explained variance is high. Table 4 shows that the two involvement components accounted for 33.10% of the explained variance. Only 10.98% of the variance is properly explained by familiarity. That is, only experience has the expected sign and it explains only 10.98%.

TABLE 3  
RESULTS OF REGRESSION ANALYSES

Beta Weights (F Statistic Below)																			
Equation Number	Knowledge	Normative Importance	Experience	Non-commitment	Sex	Single	Married	Divorced or Separated	Professional	White Collar	Blue Collar	Housewife	Student	Retired	Age	Income	Widowed	Adjusted R <sup>2</sup>	F Statistic
(1)	.157 12.54*	-.092 4.01*	.105 6.96*	.216 32.71*	.129 17.53**	.291 13.47**	.279 11.04**	.243 16.01**	-.039	-.070	.057	.082	-.003	.152	-.043	.093		.159	7.81*
(2)	-.146 11.03*	-.089 3.87	.102 6.66	.219 34.21*	.164 17.23											.065 2.70	-.145 13.94**	.127	13.44*

\* p < .05, one tailed test

\*\* p < .05, two tailed test

TABLE 4  
CONTRIBUTION OF INDEPENDENT VARIABLES TO EXPLAINED VARIANCE IN REGRESSION EQUATION (2)

Independent Variable	Percent Contribution
Knowledge	15.69
Normative Importance	9.56
Experience	10.98
Commitment	23.54
Sex	17.63
Income	6.99
Widowed	15.59
	100.00

Implications

This study suggests that an involvement-based product, really product-in-consumption situation, classification has more than face validity. Across the products used in the research, consumers could generally be classified as more or less active in their acquisition behavior. Further, using correlation-based methods, involvement was shown to be strongly related to acquisition behavior.

The study should also serve as a reminder that consumer acquisition of low involvement products is done without, the commonly assumed, meticulous examination of available brands. Despite the efforts of marketers to differentiate their brands, the lack of commitment suggests that consumers perceive brands in low involvement classes as near perfect substitutes.

An interesting question arises in considering how consumers arrived at the stage of less active acquisition behavior with low involvement products. The traditional consumer behavior models suggest active consumers directly engaging in firsthand extensive and limited problem solving behavior at earlier points in time. The logic of low involvement theory offers the alternative explanation

of a more passive consumer who at the point of first purchase is content to rely upon product information that was indirectly "caught" and not directly sought out. Information could be caught in several ways. This could include spectator-like observation of the prior purchase experience of other consumers as well as that information incidentally picked up from repetitive advertising. Such information "catching" is an alternative to the more commonly assumed information seeking.

This study also calls attention to the need to seriously consider an involvement product classification system on a segment-by-segment basis. The large residual variance in the ANOVA reported in the study suggests regularity of behavior, not so much at total market level, but rather less aggregated level. To the degree that market segment differences in product involvement can be systematically understood, the classification system offers more promise.

An immediate implication is that no one classification device is a very satisfactory device for totally accounting for differences in acquisition behavior. Explained variance is low. If product classification systems are to be of real value they must be multifaceted. For example, combination of a market structure approach to product classification (e.g., Aspinwall, 1961) and a behavioral system, similar to what is examined here would seem to provide a useful point for future research. Since any one theoretical approach is lacking, the low involvement theorists must be aware of different explanations of the same phenomena.

Appendix

Two consumers, H and M, live in the same neighborhood in Philadelphia. Both purchase the same products, but they make up their minds about what to buy in very different ways. Depending on what is being purchased, you may be like either Consumer H or Consumer M.

Consumer H

When going shopping, Consumer H feels that little knowledge can be drawn from previous shopping. Consumer H feels a lack of knowledge about the different brands or makes and models that are available in the stores. H is uncertain as to what features ought to be used in evaluating the different brands.

As a result, Consumer H actively looks for information about different brands as well as ways of evaluating the brands. Very often Consumer H will go out of the way to get information from trusted friends. Consumer H may also look at advertising for information.

Consumer H actively searches for information about the products on an extensive basis. H takes the time to learn about the products. Consumer H shops and compares the different brands.

#### Consumer M

Consumer M has the feeling of extensive knowledge when making a purchase. M knows what the different brands have to offer and, in fact, has fairly strong likes and dislikes for a few brands.

As a result, Consumer M buys with little, if any, real thought. Consumer M does little searching for information before making a purchase. M typically buys without having made comparisons between alternative brands. Consumer M feels that past purchasing experience can be relied upon.

#### References

- Aspinwall, Leo V. Four Marketing theories. Boulder: University of Colorado, 1961, chapters 3-4.
- Belk, Russell W. "Situational Variables and Consumer Behavior," Journal of Consumer Research, 2 (December 1975) 157-164.
- Belk, Russell W. "It's the thought that counts: a signed digraph analysis of gift giving," Journal of Consumer Research, 2 (December 1976) 155-162.
- Hempel, D.J. "Search behavior and information utilization in the home buying process," In P.R. McDonald, ed., Marketing Involvement in Society and the Economy. Chicago: AMA, 1969, 241-249.
- Howard, John A. and Jagdish N. Sheth, The Theory of Buyer Behavior. New York: John Wiley, 1969.
- Hupfer, Nancy T. and David M. Gardner, "Differential Involvement with Products and Issues: An Exploratory study," In D.M. Gardner, ed., Proceedings, Association for Consumer Research, College Park: ACR, 1971, 262-269.
- Jolson, Marvin A. "How to double or triple mail-survey response rates," Journal of Marketing, 41 (October 1977) 78-81.
- Kassarjian, Harold H. "Presidential Address," In H. Keith Hunt, ed., Advances in Consumer Research, Vol. 5, Ann Arbor: Association for Consumer Research, 1978.
- Kassarjian, Harold H. and Waltraud M. Kassarjian, "Attitudes under low involvement conditions," In John C. Maloney and Bernard Silverman, eds., Attitude Research Plays for High Stakes, Chicago: American Marketing Association, 1978.
- Kiesler, C.A. The Psychology of Commitment. New York: Academic, 1971.
- Kirk, Roger E. Experimental Design: Procedures for the Social Sciences. Belmont: Wadsworth, 1968.
- Krugman, Herbert E. "The impact of television advertising: learning without involvement." Public Opinion Quarterly, 29 (Fall 1965), 349-56.
- Lastovicka, John L. and David M. Gardner, "Components of Involvement," In John C. Maloney and Bernard Silverman, eds., Attitude Research Plays for High Stakes, Chicago: American Marketing Association, 1978.
- Newman, J.W. and R. Staelin, "Prepurchase information seeking for new cars and major household appliances," Journal of Marketing Research 9 (August 1972) 249-257.
- Rao, P. and R.L. Miller, Applied Econometrics. Belmont: Wadsworth, 1971.
- 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 Communications Research, Vol. II, Beverly Hills: Sage, 1973.
- Robertson, T.S., "Low Commitment Consumer Behavior," Journal of Advertising Research, 16 (April 1976) 19-27.
- Rokeach, M. Beliefs, Attitudes, and Values. San Francisco: Jossey-Bass, 1975.
- Rokeach, M. "Role of Values in Public Opinion Research," Public Opinion Quarterly, 32 (Winter 1968) 547-60.
- Skinner, B.F. Beyond Freedom and Dignity. New York: Knopf, 1972.
- Sudman, Seymour, Applied Sampling. New York: Academic, 1976.
- Winer, B.J. Statistical Principles in Experimental Design. New York: McGraw Hill, 1971.

AN EXAMINATION OF EGO-INVOLVEMENT AS A  
MODIFIER OF ATTITUDE CHANGES CAUSED FROM PRODUCT TESTING

Larry M. Newman (student), The Pennsylvania State University  
Ira J. Dolich, University of Nebraska-Lincoln

Abstract

Ego-involvement as a modifier of attitudes has recently become a topic of great interest to marketers. An experiment was conducted to measure levels of ego-involvement using Sherif's Social Judgement Theory and the effects of an automobile demonstration as a communication to create changes in attitudes. The results demonstrate the usefulness of this theoretical perspective as a schema for examining communication effectiveness.

Introduction

The concept of an individual's involvement or commitment to a product class or brand has been an expanding area of interest to marketers. Beginning with the early work of Herbert Krugman (1965, 1966, 1971), there has been speculation in this literature about the relationship between level of involvement and an individual's processing of persuasive communications. However, little research exists investigating attitude changes caused by persuasive communications and the moderating effects of involvement. The present study was conceived to examine the relationship of involvement to the amount and direction of attitude change resulting from a demonstration in a Ford Pinto.

Models to evaluate attitude change caused by communications have been proposed by Osgood and Tannebaum, Festinger, and Sherif. The Congruity Model of Osgood and Tannebaum recognizes the position of stimulus (positive or negative) and the position of the receiver (pro or con) which affect evaluation of attitude in the context of the individual's frame of reference (Zajonc, 1960). However, their model does not address itself to how involvement effects the initial position of the attitude. And, Festinger's Dissonance Theory is directed to the nature of dissonance and not attitude adjustment mechanisms. Sherif's model (Sherif and Sherif, 1969) explains the direction and amount of attitude change (assimilation/contrast) based on three factors: level of ego-involvement in an attitude (high or low); direction of attitude held (pro or con); and nature of the stimulus (positive or negative). Their model was selected for this study because of its advantages over Congruity and Dissonance models.

Application of Sherif's Model

Sherif's "Social Judgment Involvement Approach" (Sherif, 1969) has primarily been used in the measurement of attitude change to communications concerning social issues. However, it is used here to evaluate attitude change toward a product. Specifically, it has been adapted to analyze direction and extent of attitude change resulting from the positive communication of an automobile demonstration. The powerful "attention getting, desire motivating" quality of the demonstration makes it an effective tool for changing attitudes (Tilman and Kirkpatrick, 1968).

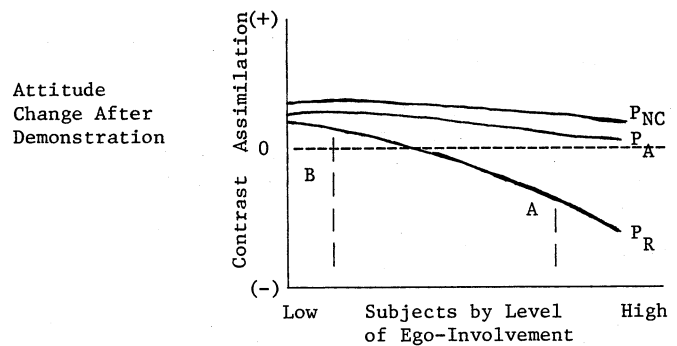
Ego-involvement is measured by the number of items the respondent accepts (latitude of acceptance) versus those he rejects (latitude of rejection) or those toward which he is neutral (latitude of noncommitment). In this study, mini-compact cars are the involvement items. Where the car being demonstrated falls within the

latitude of acceptance, there is a strong likelihood of an assimilation effect. Also, the communication will appear nearer to the subjects own position than it actually is, and thus cause a positive shift in attitude. A contrast effect will occur if the car being demonstrated falls into the latitude of rejection. The communication will be perceived as farther away from the person's own stand than it really is, thus causing a negative shift in attitude (Sherif and Sherif, 1969). Furthermore, the degree of assimilation or contrast effect is dependent upon the level of ego-involvement.

As an example, consider a hypothetical Volkswagen owner (A) who is highly ego-involved with his car and has placed the Pinto in his latitude of rejection ( $P_R$ ). The Sherif model would predict that this subject would react negatively to the communication (positive car demonstration in a Pinto) causing an attitude shift farther away from his previous position and thus deem the Pinto as more unacceptable (contrast effect).

Conversely, he might be a VW owner (B) who is low ego-involved and has placed the Pinto in his latitude of rejection ( $P_R$ ). After a demonstration ride in the Pinto (a positive communication), it is likely that he would see the merits of the Pinto and now view it closely to those cars which he accepts (assimilation effect). The theoretical relationship of these expected shifts are presented in Figure 1.

FIGURE 1  
AN EGO-INVOLVEMENT AND ATTITUDE CHANGE MODEL-THEORETICAL



Research Method

Design

A 2 x 3 (two levels of ego-involvement x three initial attitudes) between subject design was used. Subjects were self selected into the appropriate cell based on the results of an ego-involvement test. The subjects were classified as either low or high ego-involved. Their initial direction of attitude toward the Pinto was either within the region of acceptance, rejection, or noncommitment.

Subjects

Fifty university students were selected for the study. The sample was divided among those who owned mini-

compact cars and potential buyers of this size car. Car owners were randomly selected from the University's files of registered cars. Potential respondents were seniors and graduate students considering the purchase of a mini-compact car within the next year. Subjects were randomly selected from a list of senior and graduate students. These particular students were selected because they contained the highest proportion of potential new car buyers in comparison to other students. Potential mini-compact car buyers were determined by having the subjects admit it themselves. In no case did either the car owners or potential car owners have prior knowledge of the type of car they would be test driving. All subjects understood they would receive no monetary compensation.

#### Operationalization of the Independent Variable

Use of the Sherif model required a measure of ego-involvement. Photographs of twelve mini-compact cars were presented to the subjects. Each was taken from a then current issue of an automobile magazine. The photos were placed on white cardboard backgrounds with the car's name printed under the picture. Twelve cars were used to provide sufficient differences in establishing levels of ego-involvement.

An adaptation of the "Ordered Alternative Method" was used (Sherif and Sherif, 1969). After the subject studied the pictures he was put into a buyer's position by being asked to sort cars he would seriously consider buying if he had a chance to do so tomorrow. Next he was asked to sort those cars he definitely would not consider buying. The remaining cars were placed in a third pile. The cars the subjects would buy were operationally defined as the acceptance region ( $P_A$ ), the cars he would not buy his rejection region ( $P_R$ ), and the remainder the noncommitment region ( $P_{NC}$ ).

Level of ego-involvement is defined by the size of the latitude of acceptance, rejection, and noncommitment regions. Sherif et. al. (1969) state that as ego involvement increases, the size of the latitude of rejection increases compared to the latitude of acceptance. It was arbitrarily decided that those subjects with their latitude of rejection smaller than their latitude of acceptance would be considered low involvement subjects. All other subjects were placed in the high involvement category.

#### Experimental Procedure

The subjects selected for study reported to a central campus location for their individual sessions. Upon arrival, the subjects were seated and asked to complete the pre-questionnaire which consisted of identical semantic differential instruments for an ideal mini-compact car, the Pinto, and five additional mini-compact cars. The pre-questionnaire required 10 to 15 minutes to complete. As soon as the subject finished he was presented the pictures for the ego-involvement test and given instructions for accomplishing the sorts which took about five minutes. The subject was next introduced to another experimenter who presented positive stimuli about the Pinto and accompanied him on the demonstration drive. The test course was prearranged and standardized for all subjects. Total time for the drive ranged from 15 to 20 minutes. Every effort was made to disassociate the car from the experimenters thus limiting experimental demand effects. A series of statements were made to accomplish this task. A third experimenter administered the post questionnaire to the subject immediately after the demonstration drive. It was described to be similar to the pre-questionnaire. Ten minutes was usually sufficient to complete the post-demonstration drive questionnaire. Finally, the subjects were debriefed and dismissed.

#### Operationalization of the Dependent Variable: Attitude Change

A semantic differential instrument was used to determine the subject's attitude change toward the Pinto. The semantic differential was composed of 24 adjectival bipolar pairs. The scales used were mostly from Osgood's Thesaurus list (Osgood, Suci and Tannebaum, 1957), although some adaptation was required. The preliminary step in determining specific scales was taken from a tested mini-compact study (Ciotti, 1970) with the elimination of scales that showed no correlation with mini-compacts. Added in place were scales consisting of words considered to be the specific characteristics of mini-compact cars.

Individual scales were randomly assigned positions on the instrument. Then, adjectives were randomly positioned to eliminate constant scale polarity and reduce systematic bias. A similar set of scales was used for each car measured.

The questionnaires were randomly constructed by ordering the cars differently. This procedure was used to eliminate any order bias due to concept position.

The positions between each bi-polar pair were numbered from 1 to 7. The lower end of the scale denoted the position of the positive adjective. Attitude change of the Pinto was calculated by computing the differences of identical scales on the pre-demonstration questionnaire and the post demonstration questionnaire for each individual. The magnitude of the difference indicated the amount of attitude change. If the difference was a positive number, this indicated an assimilation effect. Alternately, a negative difference indicates a contrast effect.

#### Results, Discussion, and Conclusions

##### Results

The results are presented graphically in Figure 2 and summarized in Table 1. Comparison of Figures 1 and 2 indicate study results highly similar to the theoretical propositions. However, in this study, there has been a reversal of effects for individuals with the Pinto in the acceptable range with those for the Pinto in the rejection range. In addition, as compared to the theoretical propositions, there is a strong negative effect throughout.

FIGURE 2  
EGO-INVOLVEMENT AND ATTITUDE CHANGE-EMPIRICAL RESULTS

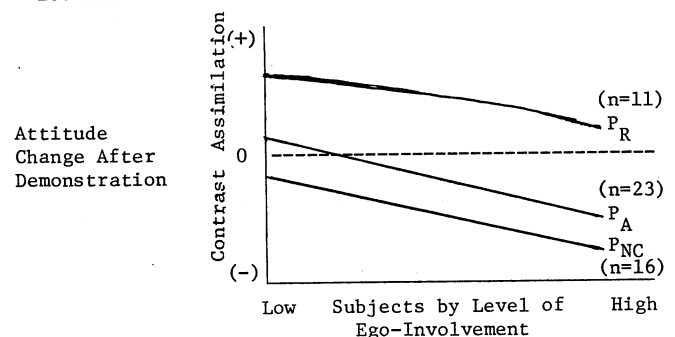


TABLE 1  
SUMMARY OF ASSIMILATION AND CONTRAST EFFECTS

	Subjects who place Pinto in their Acceptance Region.	Subjects who place Pinto in their Noncommitment Region.	Subjects who place Pinto in their Rejection region.
Low Ego-Involved Subjects	No Cumulative Change (.044)	No Cumulative Change (-.017)	Strong Assimilation Effect (.302)
High Ego-Involved Subjects	Contrast Effect (-.197)	Contrast Effect (-.252)	Moderate Assimilation Effect (+.196)

For individuals who initially place the Pinto in their acceptable range ( $P_A$ ), and are low ego-involved in the product class, no significant change occurred (although slightly positive in direction). A relatively large contrast effect did occur for highly involved individuals with the Pinto initially in their acceptable range.

For individuals who initially placed the Pinto in their noncommitment range ( $P_{NC}$ ), no significant change (although slightly negative) occurred for low involvement individuals. However, a large contrast effect did occur for highly ego-involved individuals.

The data indicates a strong assimilation effect for low involved individuals who placed the Pinto in the rejection region ( $P_R$ ). A moderate assimilation effect was obtained for highly involved individuals.

The pre-demonstration semantic differential instruments show that individuals in the acceptance region provided the most positive attitudes toward the Pinto, followed next by those in the noncommitted and finally those in the rejection region. This same relationship held for the post demonstration instruments.

#### Discussion

The results are nearly a mirror image of what would be expected from a communication perceived as a positive stimuli. However, they are generally more negative (less assimilation and more contrast) than would have been predicted. This leads to the belief that the car demonstration may have actually been a negative stimuli. Evaluation of data from Pinto owners who were used as a control group gives some credence to the negative stimuli hypothesis. Assuming that riding in another Pinto would not significantly affect their attitude toward the Pinto, control subjects (Pinto owners) were tested. Results show a slight negative shift in attitude after the demonstration drive for these Pinto owners.

The Sherif Social Judgment Theory explains attitude change based on three factors: level of ego-involvement in an attitude, direction of attitude held, and nature of the stimulus. The nature of the stimulus may be positive or negative. The expectations initially stated were on the assumption of a positive stimuli. If the communication was considered a negative stimuli by these subjects, the results would be consistent with the theory.

The findings can be interpreted in the following manner:

First, individuals have few connections with the object and do not view the demonstration as a negative experience if they are in the acceptance range with low ego-involvement. In other words, the demonstration provided acceptable congruent information. The demonstration did not have either a significant positive nor negative effect.

Secondly, for highly involved individuals, the message was discrepant to their prior attitudes. And, the information was discrepant enough to cause attitude change in a negative direction without being extreme enough in its content to cause the individual to reject the message and move the Pinto to an even more favorable position.

Finally, low ego-involved individuals in the rejection region are the most susceptible to change. The demonstration drive, although not a highly positive communication, was still positive to an individual initially rejecting the Pinto. Actually, it was a quite believable communication, thus causing a large positive shift in attitude. In the case of the highly involved, the assimilation effect occurred due to the somewhat acceptable nature of the communication. The change was not as great as in the case of the low involvement individual.

Therefore, although the results were not as originally expected, they nevertheless support the theoretical base of the Social Judgment Theory when the communication used is not strongly positive.

#### Conclusions

Marketers consider brand preferences to be based on differential attitudes and their communications attempt to alter these attitudes. Ego-involvement with the product is considered to be a major factor influencing attitude change. Although founded in Social Psychology, Sherif's Social Judgment Involvement Approach is applicable and has been adopted for this study to interpret attitude change based on a specialized type of communication.

It has often been assumed by the seller that if he can place his product before the eyes and ears of the consumer, the product will succeed or fail on its own merits. This attitude is often taken by the car salesman who assumes that the car will sell itself when driven by the prospect. This study shows that the car demonstration, like other forms of communication, is subject to assimilation and contrast effects as indicated in Sherif's model. Theory implies, as do our results, that demonstration drives can be harmful in some instances.

Future research in this area should concentrate on comparing the effects of positive and negative communications. Also, a test drive is a very unique form of communication, and while highly involving, it has many atypical characteristics. Therefore, other forms of persuasive communication could be utilized.

The Sherif's Social Judgment Involvement Approach was shown to be relevant to attitude changes created from automobile demonstrations. The study highlights problems in assuming that attempts to create positive communications from demonstrations are always favorable. Ego-involvement, as defined by the Social Judgment



Approach, was shown to have important ramifications.

#### References

Alfred P. Ciotti, "Pinto Project," unpublished paper, The Pennsylvania State University, 1970.

Herbert E. Krugman, "The Impact of Television Advertising: Learning Without Involvement," Public Opinion Quarterly, 29 (Fall, 1965), 349-356.

Herbert E. Krugman, "The Measurement of Advertising Involvement," Public Opinion Quarterly, 30 (Winter, 1966), 583-596.

Herbert E. Krugman, "Brain Wave Measurements of Advertising Involvement," Journal of Advertising Research, 11 (February, 1971), 7-9.

Charles E. Osgood, George J. Suci, and Percy H. Tannebaum, Measurement of Meaning. (Urbana, Illinois: University of Illinois Press, 1957).

Muzafer Sherif and Carolyn W. Sherif, Social Psychology (New York: Harper and Row Publishers, Inc., 1969).

R. Tilman and C. A. Kirkpatrick, Promotion: Persuasive Communication in Marketing (Homewood, Illinois: Richard D. Irwin, Inc., 1968).

R. B. Zajonc, "The Concepts of Balance, Congruity, and Dissonance," The Public Opinion Quarterly, 163 (Summer, 1960), 280-296.

## EXPLORING GENERALIZED AND PERSONALIZED BELIEFS AMONG SMOKERS AND NON-SMOKERS: A FIRST LOOK

Kjell Grønhaug, The Norwegian School of Economics and Business Administration  
Norman Kangun, The Norwegian School of Economics and Business Administration

### Abstract

In this paper we indent to look at awareness of public statements about the health hazards associated with smoking. We also want to determine what general beliefs, both health and non-health related, are held by the public, and to what extent these beliefs are personally held. The findings based on a Norwegian sample, showed great variations in generalized and personalized beliefs among smokers and non-smokers.

### Introduction

Smoking has been under attack almost continuously since tobacco was introduced to the civilized world over 400 years ago. As medical research tightened the statistical connection between incidence of smoking and various diseases (i.e., lung and throat cancer and various respiratory ailments), demands have been made in the last 15 years by various health interests groups, that governmental action should be taken to reduce smoking. To reduce the incidence of smoking, governmental bodies in both the United States and Europe have initiated a wide range of actions. Moremost among these actions has been taxation and regulation, which have been directed at influencing the conduct of producers and intermediaries. In addition, government sponsored anti-smoking information campaigns have been directed toward present and prospective smokers in the hope of discouraging smoking behavior.

This paper focuses in a limited manner on the effectiveness of the anti-smoking information program of the Norwegian government. More specifically, we seek to determine as a result of that government's campaign the extent to which Norwegian consumers are aware of public statements about the health hazards associated with smoking. We also want to determine what general beliefs, both health and non-health related, Norwegians hold with respect to the consequences of smoking, and to what extent, these beliefs are personally held. Before examining these questions, a brief discussion of anti-smoking information activities by most public authorities will be undertaken to be followed by an attempt to characterize the government model of smoking behavior which underlies their influence attempts. Furthermore, the kind of information public officials need to assess, the effects of their anti-smoking campaigns is presented.

### Governmental Anti-Smoking Information Activities

The most common information activities utilized by governmental officials to discourage smoking have been (1) advertising bans, (2) health warnings and (3) educational campaigns. Each has been used widely. (Kangun and Grønhaug, 1978.) Bans on advertising place restrictions on the conduct of sellers. The rationale for such bans is principally that the absence of positive information about smoking may reduce the propensity to smoke, particularly among young people. To date, it appears that such bans have had limited influence on cigarette consumption (Hamilton, 1977; McGuinness and Cowling, 1975). Health warnings represent a form of negative point-of-purchase information. Their use seems to be based on the assumption that continually reminding people of the health hazards associated with smoking will discourage them from smoking. The underlying idea behind the use

of educational campaigns appears to be that since many individuals lack information about the consequences of smoking, that if such information is provided intensively and continually, this may affect their smoking behavior. Actual assessment of the effectiveness of such programs has been extremely difficult.

Every program designed to affect behavior makes some assumption about how and why people behave as they do. It appears that governmental activity to influence smoking behavior, i.e., to convert smokers to non-smokers, and/or to prevent non-smokers, particularly young people, from becoming smokers, is based on the following paradigm (Kuhn, 1962; March and Olsen, 1976):

- 1) People are all alike. In this case, the underlying assumption is that everyone holds the same motives and values for smoking and will react in the same way to a given stimulus (message).
- 2) People lack appropriate information (i.e., health related information). By giving people this information, they will be properly informed, and<sup>3</sup> thus, behave as the new information directs them.

A corrolary to point 2 is as follows:

- 2a) The implicit assumption which may be deduced from most anti-smoking information programs seems to be that health related beliefs are the predominant factors for either stopping smoking or not starting to smoke. Consequently lack of information about the negative health consequences associated with smoking may be interpreted as the major reason why many people are still smoking.
- 3) Even though social aspects inherrent in smoking have been stressed in some anti-smoking campaigns, the major emphasis has been on communicating about the health consequences of smoking, thus neglecting the social aspects of smoking behavior.<sup>2</sup>

### The Kind of Information Needed to Assess the Effectiveness of Anti-Smoking Information Programs

Despite our comments in the previous section about the effectiveness of various informational programs, there is little doubt that such programs have had some effect over the last decade in reducing the incidence of smoking among adults (Fishbein, 1977). It is, however, difficult to isolate the precise effect of such activity since reduction of smoking may be attributed to other factors as well, i.e., increasing concern with physical fitness. Furing this same time span, ironically, the

<sup>1</sup>This is very similar to the "hierarchy-of-effect-models" published extensively in the marketing literature. See, among others, Lavidge and Steiner, (1961).

<sup>2</sup>This partly may be attributed to the resource situation which naturally affects the activities performed and the themes to be spotlighted. We believe that the chosen activities and themes reflect the models or paradigms held by the government officials (Sheth & Mammana, 1976).

incidence of smoking among youngsters, particularly teenage girls, appears to have increased in the past decade despite the anti-smoking activities of government (Kangun and Grønhaug, 1978).

Perhaps, the greatest limitation to assessing the impact of anti-smoking campaigns stems from lack of understanding of the smoking decision process. As noted by Fishbein (1977), there exists little systematic theory of smoking behavior despite the vastness of the literature. More importantly, we know very little about the levels of understanding among the population, regarding both health and non-health related beliefs they hold with respect to smoking. Fishbein suggests three levels of information corresponding to three levels of beliefs, that may be germane in understanding the smoking decision process. They are as follows:

- level 1 (awareness): A person may believe that "The Surgeon General or some other public agency has determined that cigarette smoking is dangerous to health".
- level 2 (general acceptance): A person may believe that "Cigarette smoking is dangerous to health".
- level 3 (personalized acceptance): A person may believe that "My cigarette smoking is dangerous to my health".

In this review, Fishbein concludes that relatively little is known in the United States about the public's beliefs about smoking, particularly their personal beliefs, and that there is sufficient evidence to conclude that the American public is not well informed at the present time. A review of literature and empirical data from other European countries seems to show that the situation is the same in those countries (NU 1975). In sum, we know little about the accuracy of the public's awareness of governmental statements about the health consequences of smoking. Nor do we know much about the extent to which people have accepted the health consequences of smoking on a personalized level. In the remaining sections of this paper, we shall try to shed some light on these questions based on a limited and exploratory Norwegian study.

#### Methodology

To get a better understanding of people's beliefs with respect to smoking, an exploratory study was undertaken. Lack of secondary data forced us to conduct primary research. The data were gathered by means of structured personal interviews conducted in March, 1978.

Based on national data (Statistisk Sentralbyrå, 1976) the universe was defined to include all individuals 15 years and older. Because of financial and technical restrictions, however, this universe was restricted to residents in Bergen, the second largest city in Norway. In the absence of a register of the actual population, the following sampling procedure was applied. Based on a voting register of households in Bergen, 40 addresses were selected at random. Interviewers were instructed to contact five households in a specific order around each address, and interview the first household member 15 years or older that they came across. The age screening (if in doubt) was conducted by posing the following question: "How old are you?" If the respondent was less than 15 years old, the interviewer asked: "May I talk to one of your parents or another member of the household at least 15 years old?" Interviews were conducted with 136 persons. The absence of 64 interviews from the planned 200 (40 · 5 - 64 = 136) can be attributed to either no household around the selected address, appropriate respondent not at home during interviewer

call and subsequent recall, or respondent unwilling to participate in the study. Five questionnaires also had to be deleted because of inadequate information. Thus, 131 respondents formed the basis of our sample.

#### Measurements

The variable and their measurements to be reported on were smoking behavior, motives for smoking or not smoking, awareness of public statements of the health hazards associated with smoking, general beliefs, and personal beliefs.

Smoking behavior included actual smoking/non-smoking, quantity smoked, past smoking history, and future smoking intentions. Questions<sup>3</sup> used to map this behavior are shown in Figure 1.

FIGURE 1  
Measurement of Smoking Behavior

- |         |   |     |    |      |
|---------|---|-----|----|------|
| Q 1.    | "Do you smoke"?   | Yes | No | → Q6 |
|         |   | ↓   |    |      |
| * Q 2.  | "Approximately how much do you smoke per day?"<br>(Cigarettes, cigars, pipes).  |     |    |      |
| * Q 3.  | "Have you ever considered stopping smoking"?  | Yes | No | → Q9 |
|         |   | ↓   |    |      |
| * Q 4.  | "Did you consider this seriously"?  | Yes | No | → Q9 |
|         |   | ↓   |    |      |
| * Q 5.  | "Did you succeed in stopping to smoke"?   | Yes | No | → Q9 |
|         |   | ↓   |    |      |
|         | "How long did you stop"?  |     |    | → Q9 |
| ** Q 6. | "Have you ever considered smoking"?   | Yes | No | → Q9 |
|         |   | ↓   |    |      |
| ** Q 7. | "Have you ever tried smoking"?  | Yes | No | → Q9 |
|         |   | ↓   |    |      |
| ** Q 8. | "How long did you smoke"?   |     |    |      |
| Q 9.    | Try to think 12 months from now. Do you think you will be smoking?  |     |    |      |
| * Q 10. | Compared to how much you smoke today, do you think (in twelve months) you will smoke less/more/or about the same per day? |     |    |      |
- \* Smokers only  
\*\* Non-smokers only

Previous research has uncovered several motives that may be relevant to either smoking or non-smoking behavior. Based on the research of Leventhal & Avis (1976), and Jaccard (1975), and related studies by NU (1975), National Council (1977), and Sosialdepartementet (1969), the following general motives have been identified as being information in the smoking decision process - health, habit, social reward, pleasure and stimulation. These motives were utilized in this study. For each of the more specific components of these general (see Figure 2), respondents were asked to rate their perceived importance (5 = extremely important, ..., 1 = not at all important).

<sup>3</sup>In order to make comparisons, these measurements are somewhat similar to what has been used in previous research (cf. National Council on Smoking and Health 1977).

FIGURE 2  
Motives related to smoking behavior

Health-Related Factors

- To take care of health
- To avoid bronchitis
- To avoid heart problems
- To be in good physical condition
- To avoid cancer (throat)
- To avoid ulcer

Habitual Factors

- To possess fixed habits
- To have something in the mouth

Social Rewards

- To feel accepted among strangers
- To be accepted among friends

Pleasures

- To experience pleasure
- Not to experience unpleasantness

Stimulation/Well-Being Factors

- To be relaxed
- Not to feel "blue"
- To keep calm
- To feel energetic

To determine the accuracy and awareness of who said what about smoking and its consequences, respondents were shown the following statements:

1. Smoking may be dangerous to your health
2. If one stops smoking, one will usually gain weight
3. The incidence of heart attacks increases if one smokes
4. Physical fitness is inconsistent with smoking.
5. Smoking may cause lung and throat cancer
6. Smoking may cause bronchitis
7. Smoking can lead to the development of ulcers.

After being shown each statement, the respondents then were asked to identify the agency making the statement. They were given a list of four agencies which are shown here:

- (1) The Norwegian Tobacco Council
- (2) The Consumer Ombudsman
- (3) The Norwegian Medical Association
- (4) National Council on Smoking and Health

To ascertain generalized beliefs (level 2), the respondents were given a series of statements and then asked to indicate the extent to which they agreed with these statements. Some of the statements used were as follows: It is pleasant to smoke; It is relaxing to smoke; and Smoking may cause bronchitis. With respect to identifying personal beliefs (level 3), the respondents were confronted with the statements similar in content to statements above, but phrased in a personal way: I find pleasure in smoking a cigarette; I smoke to relax; and If I smoke, I may get bronchitis. For each of these statements, respondents were asked the extent to which they agreed or disagreed (5 = strongly agree, . . . , 1 = strongly disagree).

Finally some socioeconomic and demographic data was collected. Respondents were classified according to sex, age, education, and household income.

Findings

The sample consisted of 131 respondents - 70 (i.e., 53%) were smokers and 61 (i.e., 47%) were non-smokers. The proportion of smokers in this sample is somewhat higher than that found in recent national surveys (cf. National Council on Smoking and Health, 1977). With respect to age, the highest percentage of smokers were found in the 25-35 age bracket. Men were found to be more frequent smokers than women, and the incidence of smoking was found to be inversely related to the level of education attainment. All of these results are similar to findings in recent Norwegian and U.S. surveys.

When looking at the smoker sub-sample, 55 of 70 (79%) claimed that they had tried to stop smoking, and 50 of these respondents (71%) maintained that they had made serious attempts. Approximately half of those who really tried to stop smoking, claimed that they succeeded in doing so, but for the majority, this period of abstinence lasted less than three months. Of the non-smokers sub-sample, 45 of 61 (74%) had previously tried smoking. The time period for which they had smoked was denoted by a U-shaped curve indicating a significant number of both less than two-week smokers, and more than two-year smokers. These findings indicate that some individuals are capable of trying smoking without becoming permanent adopters of that behavior while some adopters are able to change their behavior.

When smokers were asked about their smoking behavior in the next 12 months, 21 (30%) claimed that they still would be smoking, 15 (21%) said they would not be smoking, and the remaining (49%) claimed they "did not know".

Compared to American findings reported by Fishbein (1977), these results reveal dramatic changes. At least in Norway, these results indicate that the militant anti-smoking policy of the government (Bjartveit 1977) has resulted in smokers giving serious consideration toward quitting. Furthermore, many smokers indicated they would reduce their level of consumption in the future. No one from the non-smoking group believed they would be smoking in the next 12 months. No significant differences were found across sex, age, educational, or income groups with regard to anticipated changes in smoking behavior.

Motives

Distributions of the various motives for smoking or not smoking are shown in Table 1. The table reveals that reasons for smoking or not smoking do vary in importance across individuals within the same group. Variations in perceived importance are similarly observed with respect to health related motives.

When looking at the various groups of generalized motives, high scores (i.e., high perceived importance) are found - as expected for the health related motives as well as for "pleasure" and "stimulation/well-being".

When looking at the importance attached to certain motives by non-smokers and smokers, some dramatic differences emerge: Regarding health related motives, non-smokers' scores for all six motives in this category are higher than smokers scores. If we assume independence between the various health related motives, the probability of getting such a result is  $P(B=6) = .0156$ ,  $n=6$ ,  $p = \frac{1}{2}$ . When examining the proportion of respondents in each group that assigned extreme values, significant differences occur for "taking care of health" ( $p < .01$ ); "avoiding bronchitis" ( $p < .001$ ), "avoiding heart problems" ( $p < .05$ ), and "being in good physical condition" ( $p < .05$ ).<sup>4</sup>

<sup>4</sup>The differences among proportions have been tested by using t-test (cf. Blalock 197, p. 228-232). The tests are two-tailed, because of the absence of explicit hypotheses.

TABLE 1  
Motives for Smoking and Not Smoking and Their Perceived Importance (%)

Motive	Perceived Importance:						Total
	Non-smokers N=61			Smokers N=70			
	H	M	L	H	M	L	
<u>Health</u>							
1. Take care of health	88	7	5	68	20	12	100
2. Avoid bronchitis	81	10	9	54	23	23	100
3. Avoid health problems	85	8	7	77	13	10	100
4. Good physical shape	80	15	5	62	29	9	100
5. Avoid cancer	86	9	5	78	16	6	100
6. Avoid ulcer	84	11	5	79	15	6	100
<u>Habit</u>							
7. Fixed habits	40	15	45	20	31	49	100
8. Have smoking in the mouth	5	3	92	13	10	77	100
<u>Social reward</u>							
9. Accepted by strangers	66	17	17	48	29	23	100
10. Accepted among friends	54	31	15	39	29	32	100
<u>Pleasure</u>							
11. Experience pleasure	73	18	8	59	31	10	100
12. Unpleasant	62	18	20	56	25	19	100
13. Pleasant smell	30	25	45	25	23	52	100
<u>Stimulation/well being</u>							
14. Relaxation	72	17	11	64	27	9	100
15. To not feel "blue"	24	14	62	23	17	60	100
16. Calm	67	21	12	65	28	7	100
17. Energetic	66	29	5	75	18	7	100

For habit-related motives, non-smokers do place more emphasis on "fixed habits" than do smokers ( $p < .02$ ). However, smokers attach more importance having something in their mouths ( $p < .10$ ).

With respect to the remaining generalized motives, the following findings emerge. Non-smokers stress social reward more than smokers.

No significant differences are found for either the pleasure or stimulation/well-being motive categories. With respect to the socio-economic characteristics in this study, no significant differences were found.

Awareness

Table 2 reports on the levels of awareness among respondents in associating statements about the health hazard of smoking with the appropriate public and/or private bodies. The Consumer Ombudsman, The Norwegian Medical Association and The National Council of Smoking and Health have all declared that smoking may be dangerous to health, contribute to heart problems, cause lung and throat cancer, ulcers, bronchitis and lead to inferior physical fitness. Furthermore, the National Council of Smoking and Health has tried to inform people how to avoid putting on weight after giving up smoking. Aside from required warnings on packages, the tobacco industry has not given the public any information about the health hazards associated with smoking.

The results presented clearly indicate that a substantial portion of the public is not aware of the pronouncements by public agencies about smoking. For example, even on the statement relating smoking to lung and throat cancer, substantial proportions of both smokers and non-smokers are not aware of agency communications.

When comparing the awareness among non-smokers and smokers with respect to these statements, an interesting finding emerges. It appears that smokers are equally as aware of these statements as non-smokers. This result stands in contradiction to "dissonance theory" which

TABLE 2  
General Awareness of Public Statements About the Consequences of Smoking or Not Smoking (Level 1 Information)(%)

Statement		The Norway Tobacco Ind.			The Cons. Ombudsmann			The Norway Med. Ass.			National Cons. of Smo. & Health			Total
		Yes	No	Don't Know	Yes	No	Don't Know	Yes	No	Don't Know	Yes	No	Don't Know	
		Smoking is dangerous to health	Non-S	10	57	33	36	25	39	75	17	8	80	
	Smokers	11	75	16	34	33	33	81	13	6	87	4	8	100 (N=70)
One will gain weight by stopping smoking	Non-S	30	34	36	13	36	51	26	39	35	89	4	7	100 (N=61)
	Smokers	17	54	29	6	56	38	30	37	33	13	47	40	100 (N=70)
Smoking can increase the incidence of heart attack	Non-S	10	56	34	30	25	45	75	8	17	52	15	33	100 (N=61)
	Smokers	7	71	21	14	40	46	83	6	11	57	19	24	100 (N=70)
Smoking contributes to lack of physical fitness	Non-S	10	52	38	21	26	53	62	15	23	54	10	36	100 (N=61)
	Smokers	1	69	30	14	37	49	84	7	9	57	14	29	100 (N=70)
Smoking causes cancer (throat)	Non-S	6	56	38	25	23	52	75	12	13	57	13	30	100 (N=61)
	Smokers	-	30	70	16	41	43	83	11	6	59	14	27	100 (N=70)
Smoking causes bronchitis	Non-S	12	51	37	20	26	54	67	12	21	48	18	34	100 (N=61)
	Smokers	3	70	27	11	39	50	73	11	16	43	20	37	100 (N=70)
Smoking can cause ulcers	Non-S	15	44	41	13	31	56	53	13	34	33	23	44	100 (N=61)
	Smokers	1	70	29	10	40	50	57	10	33	36	18	46	100 (N=70)

seems to suggest that such statements should be avoided by smokers since they're contrary to smoking behavior. A number of explanations can be offered to explain this finding. First, smokers may have a higher tolerance for dissonance than non-smokers, and thus not avoid pronouncements about the health hazards associated with smoking. Second, smokers may either overlook the consequences of smoking, or tend to give health-related consequences less weight in the decision to smoke. And, finally, smokers may accept the consequences, but either distort the consequences through reinterpretation or distort the information.

TABLE 3  
General Beliefs Among Non-Smokers and Smokers (%)

General Statement About:	Non-smokers <sup>x)</sup> (N=61)			Smokers (N=70)			Total
	A	N	D	A	N	D	
<u>Health-Related Factors</u>							
1. Smoking is <u>not</u> dangerous to health	5	3	92	12	1	87	100
2. Smoking may cause bronchitis	64	25	10	51	30	19	100
3. Smokers are more inclined to have heart attack than non-smokers	83	14	3	73	21	6	100
4. Smoking does <u>not</u> contribute to an inferior physical condition	20	5	75	17	10	73	100
5. Smoking may cause cancer (throat)	81	7	12	60	26	14	100
6. Smoking cannot cause ulcers	51	43	6	19	50	31	100
<u>Habitual Factors</u>							
7. It's usual for smokers to automatically light a cigarette	49	15	36	44	10	46	100
8. The reason many people smoke is to have something in the mouth	29	24	47	10	17	73	100
<u>Social Rewards</u>							
9. When smoking, acceptance by strangers is more difficult to achieve	48	30	22	34	29	37	100
10. When smoking, one is more accepted among friends	24	22	54	11	11	77	100
<u>Pleasures</u>							
11. Smoking is pleasant	16	23	61	46	33	20	100
12. Being without cigarettes is not unpleasant	61	7	32	20	10	70	100
13. The smell of tobacco is unpleasant	42	32	26	13	38	49	100
<u>Stimulation/Well-being Factors</u>							
14. Smoking is relaxing	32	22	46	79	6	15	100
15. Cigarettes help overcome "blue" moods	29	26	45	50	20	30	100
16. Smoking has a calming effect on smokers	35	22	43	64	16	20	100
17. Smoking can make people feel better	22	23	55	44	26	30	100

x) A(gree)=5,4 N(eutral)=3, D(isagree)=2,1

## General beliefs

Table 3 reports on general beliefs (level 2 information) related to various aspects of smoking among the respondents. Some interesting variations in the degree of agreement/disagreement occur between smokers and non-smokers. With respect to general beliefs about the health consequences of smoking, a greater proportion of non-smokers seem to believe that smoking can lead to impaired health than do smokers. For example, significant differences between these two groups are found for statements 5 ( $p < .01$ ) and 6 ( $p < .01$ ). Ironically, non-smokers are more likely to express the belief that smoking cannot cause ulcers than smokers - the opposite of the expected direction.

When looking at the statements related to pleasure and stimulation/well-being, significantly more smokers express positive beliefs than do non-smokers. All the differences are significant at the .001-level. For example, significantly more smokers (46%) than non-smokers (16%) believe smoking to be pleasant. Or, almost 4/5's of the smokers agree that "smoking is relaxing" whereas less than one-third of the non-smokers agree with that statement. Interestingly, the greatest differences between these two groups lie in the pleasures-stimulation categories, yet most of the anti-smoking information campaigns are directed toward influencing health-related beliefs. Yet among smokers, these beliefs may be less important in the decision to continue to smoke than beliefs in the pleasures and stimulation areas.

When relating general beliefs to the various socio-economic characteristics, no differences were found for the general beliefs listed in Table 3.

## Personalized beliefs

Table 4 focuses on personalized beliefs related to various aspects of smoking. The interesting thing about Table 4, particularly as compared to Table 3, is that significant differences emerge between the two groups on almost all the personalized statements. Smokers generally tend to exhibit less agreement than non-smokers with respect to the consequences to their personal health of smoking. The greatest differences between these two groups in this general category are found in statements 3 ( $p < .02$ ), 5 ( $p < .01$ ) and 6 ( $p < .05$ ). Smokers also tend to agree more extensively with importance of habit in maintenance of their smoking behavior than do non-smokers. For example, significantly more smokers agree with the statement about automatically lighting a cigarette than do non-smokers ( $p < .001$ ).

The smoker and non-smoker sub-samples differ in their views on smoking and social rewards. Non-smokers were more likely to perceive the social rewards of smoking negatively than smokers, i.e., for statements 9 and 10, the differences between the two groups were  $p < .001$  and  $p < .05$  respectively. For personalized beliefs regarding pleasures, smokers, as one might expect, found smoking pleasant in various ways. All differences in these statements were significant at .001 level. Finally, smokers exhibited significantly more agreement about the personal well-being derived from smoking, i.e., relaxation, calming effect, energy, etc. than did non-smokers. The differences here were also significant at the .001 level.

In evaluating personalized beliefs, smokers who had tried to stop smoking were compared with those who had not tried to stop. The former exhibited more pleasure in smoking ( $p < .02$ ) and were more distressed to be without cigarettes ( $p < .05$ ) than the latter. They also believed that they were more accepted among friends when smoking ( $p < .05$ ), found cigarettes more helpful when "blue" ( $p < .001$ ), and more effective in calming them down ( $p < .05$ ).

TABLE 4

## Personalized Beliefs Among Non-Smokers and Smokers (%)

Information About:	Non-smokers (N=61)			Smokers (N=70)			To- tal
	A	N	D	A	N	D	
<u>Health-Related Factors</u>							
1. Smoking is not dangerous to <u>my</u> health	4	5	91	6	11	83	100
2. <u>My</u> smoking may result in <u>my</u> getting bronchitis	66	21	13	47	41	11	100
3. <u>My</u> smoking may result in <u>my</u> having a heart attack	75	11	14	63	31	6	100
4. If <u>I</u> smoke, it will not influence <u>my</u> physical condition	20	5	75	19	17	64	100
5. <u>My</u> smoking may lead to <u>my</u> getting cancer (throat)	80	13	7	57	31	11	100
6. <u>I</u> will not get ulcers if <u>I</u> smoke	20	35	45	20	53	27	100
<u>Habitual Factors</u>							
7. <u>I</u> often light a cigarette without thinking about it	7	7	86	49	11	40	100
8. <u>I</u> smoke to have something in my mouth	7	6	87	12	16	72	100
<u>Social Rewards</u>							
9. <u>I'm</u> frequently rejected by strangers when <u>I</u> smoke	71	9	20	38	20	42	100
10. When <u>I'm</u> smoking, <u>my</u> friends find me more acceptable	4	9	87	7	24	69	100
<u>Pleasures</u>							
11. <u>I</u> find personal pleasure in smoking	5	4	91	73	20	7	100
12. It is not unpleasant for <u>me</u> to be without cigarettes	88	3	9	31	11	37	100
13. <u>I</u> do not like the smell of tobacco	39	32	29	14	36	50	100
<u>Stimulations/Well-being Factors</u>							
14. <u>I</u> find that smoking is relaxing	4	5	91	73	21	6	100
15. When <u>I'm</u> feeling "blue", <u>I</u> find cigarettes helpful in overcoming these moods	5	13	82	48	29	23	100
16. Smoking calms <u>me</u> down when <u>I'm</u> tense	7	13	84	66	27	7	100
17. <u>I</u> feel better (more energetic) when <u>I'm</u> smoking	5	9	86	49	30	21	100

When examining the differences in personalized beliefs among smokers by daily consumption, heavy smokers (i.e., those smoking 20 or more cigarettes a day) felt it to be more unpleasant to be without cigarettes than did light smokers ( $p < .01$ ). No significant differences were found when relating socio-economic characteristics to the various statements of personalized beliefs.

## Differences in generalized and personalized beliefs

Fishbein (1977) assumed the negative consequences of smoking to health would be less accepted by smokers at the personalized than at the generalized level. Table 5 is based on the extreme values extracted from Tables 3 and 4. As fishbein predicted, lower values were found among smokers for all the six health-related statements at the personalized level vis-a-vis the generalized level ( $p < .02$ ). For the non-smokers, a higher acceptance of the health-related statements was observed at the personalized than at the generalized level. However, the results relating smoking to ulcers was interesting. At the generalized level only 6% of the non-smokers disagreed with the statement "smoking cannot cause ulcers", while at the personalized level 45% disagreed with this statement ( $p < .001$ ). This result may indicate that negative effects attributed to smoking by non-smokers partly based on relevant information, may also be generalized to other events where they lack information.

TABLE 5  
Differences in Generalized and Personalized Beliefs

Statement/Information About:	Non-smokers			Smokers		
	(1)	(2)	(3)	(1)	(2)	(3)
	Diff. (1)-(2)			Diff. (1)-(2)		
	Gen.	Pers.	(2)	Gen.	Pers.	(2)
<u>Health</u>						
1. Not dangerous to health <sup>xx</sup> )	92	91	1	87	83	4
2. Smoking and bronchitis <sup>x</sup> )	64	66	-2	51	47	4
3. Smoking and heart attack <sup>x</sup> )	83	75	8	73	63	10
4. Smoking does not influence physical condition <sup>xx</sup> )	75	75	0	73	64	9
5. Smoking and cancer <sup>x</sup> )	81	80	1	60	57	3
6. No relationship between smoking and ulcers <sup>xx</sup> )	6	45	-39	31	27	4
<u>Habitual Factors</u>						
7. Habit and smoking <sup>x</sup> )	49	7	42	44	49	-5
8. Have something in the mouth <sup>x</sup> )	29	7	22	10	12	-2
<u>Social Rewards</u>						
9. Not accepted by strangers <sup>x</sup> )	48	71	-23	34	38	-4
10. Acceptance among friends <sup>x</sup> )	24	4	20	11	7	5
<u>Pleasures</u>						
11. Smoking and pleasure <sup>x</sup> )	16	5	11	46	73	-27
12. Independence of tobacco <sup>x</sup> )	61	88	-27	20	31	-11
13. Unpleasant smell of tobacco <sup>x</sup> )	42	39	3	13	14	-1
<u>Stimulation/Well-being Factors</u>						
14. Relaxing <sup>x</sup> )	32	4	28	79	73	6
15. Smoking helps overcome "blue" moods <sup>x</sup> )	29	5	24	50	48	2
16. Calming effect of smoking <sup>x</sup> )	35	7	28	64	66	-2
17. Feeling better (more energetic) when smoking <sup>x</sup> )	22	5	17	44	49	-5
	(N=61)			(N=70)		
	x) A(gree)			xx) D(isagree)		

For the statements related to habitual factors and social rewards, the changes are much more dramatic for non-smokers than for the smokers.<sup>5</sup> As shown in Table 5, non-smokers, to a lesser extent, accept smoking as habitual and social rewards from smoking at the personalized level. All the differences were significant at the .001 level. For smokers, however, no systematic changes were observed between generalized and personalized beliefs.

Turning to the statements about pleasures and stimulation/well-being, remarkable decreases in the degree of acceptance at the personalized level were found among smokers. Non-smokers, however, on these same dimensions personally found smoking less pleasurable, felt less dependent on tobacco ( $p < .001$ ), found smoking less relaxing ( $p < .001$ ), less helpful in overcoming depression ( $p < .001$ ), and felt it did not have a calming effect ( $p < .001$ ) than they generally believed. In contrast, smokers found smoking more pleasurable ( $p < .001$ ) personally than they generally believed. For the other items modest, not significant, changes were observed indicating a high degree of correspondence between personalized and generalized beliefs.

#### Analysis

The findings that emerge from this exploratory study are as follows:

1. Many smokers and non-smokers exhibit uncertainty about the risks to health associated with smoking.
2. Respondents in both groups expressed doubt on both a general and personal level about public statements regarding the health hazards associated with smoking.
3. The smoker and non-smoker sub-samples differed significantly on many of the motives, generalized beliefs and personalized beliefs associated with smoking behavior.
4. The socio-economic characteristics of our sample population were found to be less useful in revealing significant differences in smoking behavior.

Several implications may be drawn from the reported findings of this exploratory study. For the public policy makers intent on curbing the incidence of smoking, changing smokers to non-smokers will require programs with broader focus. These programs should encompass factors other than health related beliefs in seeking to influence smoking behavior. Our data show that pleasure and social consequences associated with smoking, must also be taken into account and dealt with when designing anti-smoking communication strategies.

For researchers, the present study, being exploratory in nature, does pose questions which need to be investigated more thoroughly. As noted at the outset of this paper, we need to know more about people's personal beliefs about smoking and non-smoking. Since these beliefs are most material to the smoking decision, lack of information about them is critical and raises serious questions about the success of government initiatives to discourage smoking. Emphasis also should be placed on longitudinal studies in order to map changes in perceived beliefs among relevant populations. Emphasis, of course, should be placed on tracing factors which account for the differences between smokers and non-smokers. Large scale studies of specific segments of smokers and non-smokers would also be of extreme importance in allow-

ing us to gain a better understanding of this complex phenomenon.

A note of caution is appropriate in closing. Behavioral change, of any kind, is neither easy nor automatic. Knowing more about smoking may not result in changing the behavior of current and prospective smokers. Thus, we may know *why* people act as they do and be powerless to change them using morally acceptable procedures. Knowing less, however, almost certainly guarantees that most anti-smoking information programs will fail. The serious consequences from smoking, both at the individual and societal level, deserve the best of our efforts, not the least.

#### References

- Bjarveit, "The Norwegian Tobacco Act," Health Education Journal, Vol. 36, No. 1, 1977 (reprint).
- H. B. Blalock, Social Statistics (New York: McGraw-Hill Book Company, 1972)
- Fishbein, Consumer Beliefs and Behavior with Respect to Cigarette Smoking: A Critical Analysis of the Public Literature (Staff Report of the Federal Commission, May 1977)
- J. L. Hamilton, The Cigarette Advertising Ban as a Public Health Policy, Working Paper No. 5, Wayne Economic Papers (Detroit: Wayne State University, 1977)
- N. Kangun and K. Grønhaug, Government Anti-Smoking Information Campaigns: Gaps in the Strategy of Deterrence (Bergen: Paper presented at Corporate Social Concerns and Public Policy Conference, April 2-4, 1978)
- T. S. Kuehn, The Structure of Scientific Revolutions (Chicago: The University of Chicago Press, 1962)
- R. C. Lavidge and G. A. Steiner, "A Model of Predictive Measurement of Advertising," Journal of Marketing, 25 (October, 1961), 59-62.
- H. Leventhal and Nancy Avis, "Pleasure, Addiction and Habit: Factors in Verbal Rapport on Factors in Smoking Behavior," Journal of Abnormal Psychology, 85(1976), 478-488.
- J. McGuinness and K. Cowling, "Advertising and the Aggregate Demand for Cigarettes," European Economic Review, 6(1975), 311-328.
- J. G. March and J. P. Olsen, Ambiguity and Choice in Organizations (Bergen: Universitetsforlaget, 1976)
- National Council on Smoking and Health, Evaluation of Anti-Smoking Programs and Activities (Oslo: 1977)
- N.U., Nordiska Tobaksarbetsgruppens Betankande (Helsingfors: 1975)
- J. N. Sheth and N. J. Mammana, "Recent Failures in Consumer Protection," California Management Review, Vol. XVI, Spring 1974, 64-72.
- Sosialdepartementet, Om Pavirkning av Røykevaner, St. meld. nr. 62 (1968-69).
- Statistisk Sentralbyrå, Røykevaneundersøkelse 1974, Rapport fra underavdelingen for intervjuundersøkelser (Oslo: Statistisk Sentralbyrå, 1976)

<sup>5</sup>The differences of differences (cf. Blalock 1972, p. 230-232) are all significant at the .001 level.



# INVOLVEMENT: A POTENTIALLY IMPORTANT MEDIATOR OF CONSUMER BEHAVIOR

Andrew A. Mitchell, Carnegie Mellon University

## Abstract

Although "involvement" has the potential of being an important mediator of consumer behavior, our current understanding of its effects are limited. The primary reason for this seems to be our general failure to develop a publicly acceptable conceptual definition of "involvement", valid measures of it and procedures for manipulating it in the laboratory. Until this is accomplished the quality and quantity of empirical research in this area will be limited. The papers in this session, in general, reflect this problem.

## Introduction

Based on the number of papers on the subject over the last couple of years, interest in "involvement" and its effect on consumer behavior has been growing at an increasing rate (e.g. Robertson, 1976; Rothschild, 1978). There seems to be general agreement that "involvement" is a potentially important mediator of consumer behavior, however, "involvement" remains an elusive concept. Precisely what is involvement? How do we measure it? How do we manipulate involvement in the laboratory? Until we can answer these questions, the quantity and quality of empirical research on the subject will remain limited.

In this discussion paper, I would like to elaborate on these issues after briefly discussing the three papers in this session. Of the three papers, the one by Gronhaug and Kangan seems to have only a tangential relationship to "involvement." Therefore, I will discuss it first, followed by the Newman and Dolich and the Lastovicka papers, respectively. These latter two papers present the results of empirical studies using different measures of "involvement." In the discussion of these papers, I will only mention how "involvement" was measured in each study and reserve discussion of these measures until later.

## Gronhaug and Kangan Paper

This paper examines differences in smoking related beliefs between smokers and nonsmokers in Norway. Although there is some question as to the projectability of the results, the study does indicate that nonhealth related beliefs may be important determinants of smoking behavior. In addition, the overall measurement approach and the cross cultural nature of the study make it an interesting paper.

## Belief Hierarchy

In the study, the authors use Fishbein's hierarchy of beliefs or different levels of beliefs (Fishbein, 1977) to analyze smoking behavior. These different levels are awareness, general acceptance and personal acceptance. They form a hierarchy because a belief at the personal acceptance level, for instance, implies a belief at the awareness and general acceptance levels. This conceptualization of different belief levels has two important implications. First, it suggests that information at the different levels may be organized in different structures in secondary or long term memory. For instance, there may be a memory structure for the general concept of smoking which contains general beliefs about smoking. There may also be a separate memory structure containing beliefs about an individual's smoking behavior. This

conceptualization is consistent with some of our preliminary findings using elicitation procedures to examine the content of stored information about specific brands in secondary memory (see Olson, 1979 for a discussion of these procedures). Basically we find that placing individuals in a purchase mode produces a different set of associations about a particular brand than if the subject were not in this mode. As might be expected, the associations elicited from a purchase mode are generally evaluative (i.e. brand x gets clothes cleaner).

In general, we would expect that the beliefs organized into the memory structure concerning the general concept of smoking would determine a general attitude about smoking, while beliefs organized into the memory structure concerning the individual's smoking behavior would determine the attitude toward the act of smoking for the individual. Consequently, we would expect that personal beliefs would be better predictors of behavior. This study provides some evidence of this since more differences were found between smokers and nonsmokers for beliefs at the personal acceptance level than the general acceptance level.

A second implication is that different cognitive processes are required to form beliefs at the different levels. The formation of beliefs at the awareness level seems to involve processes similar to those found with incidental learning (Postman, 1964; McLaughlin, 1965). Here information is learned at the nonsense syllable level (i.e. little meaning is attached to the information). At both the general acceptance and the personal acceptance levels the information acquires semantic meaning. At the personal acceptance level, however, the information has been directly related to the individual's behavior and the outcomes of this behavior. A similarity should be noted here with Krugman's measure of "involvement" as the number of "connections" an individual makes between a message and the "content of their life" during exposure to a message (Krugman, 1967). Cognitive processes that result in the generation of these "connections" would be essential in forming beliefs at the personal acceptance level. However, these "connections" could also be made after exposure to a message. Individuals may form beliefs at the general acceptance level and then, at some later point in time, reprocess this information to form beliefs at the personal acceptance level.

In summary, the concept of different levels of beliefs is an interesting one and is applicable to other areas of consumer research. As we become more concerned with measuring precisely what information consumers store in secondary or long term memory about products, the distinction between different belief levels should prove useful. This distinction may also prove useful in communication research. It may be determined, for instance, that the type of copy strategy used in an advertisement may have an important affect on whether individuals form beliefs at the personal acceptance level.

## Criticisms

My criticisms of the paper are relatively few in number. First, in the presentation of the results, it would be helpful if there was an indication in the tables of which variables were significantly different. It is also not clear from the paper, exactly what statistical tests

were used to determine statistical significance. Second, I believe that the awareness level should be measured by a general awareness of information as opposed to knowledge of which particular agency disseminated it. Someone may remember seeing a message that "smoking causes cancer," but may not remember the sponsoring agency. In using the proposed framework, recognition of the information would seem to be a more important measure than recall of the sponsoring agency. Finally, although the paper stresses the importance of understanding the causes of smoking behavior and how it may be changed, the thrust of this research does not appear to be in this direction.

#### Future Research

In using Fishbein's attitude theory (Fishbein, 1967) to examine how smoking behavior may be changed, the first step is to examine precisely which beliefs determine smoking behavior. Here we may find considerable heterogeneity in the population with respect to these beliefs, which may form a basis for segmenting the market. Second, it must be determined if getting smokers to form beliefs about the health hazards of smoking at the personal level will change their smoking behavior. Again, we should expect to find that the formation of these beliefs may cause some smokers to change their behavior, but not others. Finally, if it is determined that getting smokers to form these beliefs will cause a significant number of smokers to change their behavior, then a communication program must be developed to accomplish this goal. This, of course, is probably the most difficult step. If it is determined that getting smokers to form beliefs about the health hazards of smoking at the personal level will not cause a change in the behavior of smokers, then, other beliefs will need to be attacked or formed.

This discussion raises one final political question that does not seem to have been resolved, at least in the United States. This concerns to what extent the government should attempt to deter cigarette smoking. At one level the government may simply be concerned with informing its citizens about the health hazards of smoking. The goal would be to have significant proportion of the population form beliefs about the health hazards of smoking at the general acceptance level. At the second level, the government would want a significant proportion of the population to form these beliefs at the personal acceptance level. At the next level, the government would actively try to deter cigarette smoking by any means possible short of banning cigarettes from the market. For instance, attempts may be made to change nonhealth related beliefs if this proves more effective in changing smoking behavior than changing or forming health related beliefs. A final level, would be the outright ban of cigarettes. I suspect, however, that the government will not want to make this decision until an understanding of the relative effectiveness of programs based on these different levels is determined. This, I think, emphasizes the importance of this research.

#### Newman and Dolich Paper

This paper examines "ego involvement" as a mediator of communication effects using Social Judgment Theory (Sherif, Sherif and Nebergall, 1965) as a framework. It is somewhat unique in that the communication was a demonstration drive in an automobile. Clearly, an understanding of how "involvement" mediates communication effects is needed, however, conceptual and methodological problems in this study place into question the validity of the results.

#### Methodological Problems

From a methodological perspective, it should be noted first that this basically is a correlational study. Consequently any differences found between groups may always be attributed to any other characteristic on which the groups

differ. Second, the use of Social Judgment Theory as a framework means that a control group cannot be used to control for possible measurement effects due to the necessity of a before treatment measure. Because of this, the resulting design is subject to demand characteristics. Although different experimenters were used for each part of the study, I suspect that this did not conceal the purpose of the study from many of the subjects. Finally there are always problems in using change scores as a dependent variable. Space does not permit a thorough discussion of this problem, however, it occurs because the magnitude of a difference score is due to both actual or treatment differences and measurement error (Cronbach and Furby, 1970). Frequently, these two effects will be systematically related as in regression effects (Lord, 1963). These problems can usually be controlled by using the after measure as the dependent variable and the before measure as a covariate. In this study, however, the use of this procedure is not possible since the before measure will be highly correlated with whether the Pinto is an acceptable or unacceptable alternative. It is difficult to assess the effect of these methodological problems on the results of the study, however, it should be noted that each one of them provides one or more alternative explanations of the results.

#### Conceptual Problems

There are also conceptual problems in the interpretation and application of Social Judgment Theory in this situation. To properly use Social Judgment Theory, the authors should have measured the respondents' prior evaluation of the Pinto, the latitudes of acceptance and rejection of messages concerning the Pinto on an evaluative scale, the respondents' perception of the position advocated by the message and the respondents' evaluation of the Pinto after receiving the message (Figure 1). In this particular situation, however, the message was a demonstration drive in a Pinto, so it would be difficult to obtain the latter two measures separately. Consequently, a confirmation/disconfirmation of expectations framework (Olson and Dover, 1979; Anderson, 1973) may be more appropriate in this situation.

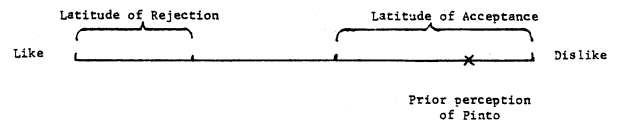


FIGURE 1  
SOCIAL JUDGMENT THEORY APPLIED TO MESSAGES EVALUATING THE PINTO

It should also be noted that the theoretical predictions given in Table 1 of the paper are not based on Social Judgment Theory. Without knowledge of the subject's latitude of acceptance and rejection for messages providing an evaluation of the Pinto and the subject's perception of the message, no prediction is possible as to either the direction or amount of attitude change. In addition, the terms assimilation and contrast are misused. In Social Judgment Theory these terms refer to distortions in perceptual judgments and not the direction of attitude change (Sherif and Hovland, 1953). Finally, ego involvement with the Pinto should be measured by the relative sizes of the latitudes of acceptance and rejection and not by the relative number of acceptable and unacceptable brands (Sherif and Sherif, 1967). This latter measure may be confounded by the location of the individual in the purchase decision process. For instance, someone just beginning the process may have a large number of acceptable brands, while someone near the end of the process may only have a small number of acceptable brands.

Aside from these methodological and conceptual comments, I also have a couple of comments concerning the presentation of the results. It would be helpful if the authors gave the number of subjects per cell, discussed precisely how attitudes were measured and mentioned what statistical tests were used, if any.

In summary, methodological and conceptual problems place into question the validity of the results. If the authors' main goal is the testing of Social Judgment Theory in a marketing context a different experimental situation will probably be required. If the authors are primarily interested in examining the effects of a demonstration drive as a method of attitude change a confirmation/disconfirmation of expectations framework may be more appropriate.

#### Lastovicka Paper

This paper examines "involvement" as a basis for a product classification system. The author suggests that in order for "involvement" to form such a basis, it must be first demonstrated that "involvement" is an important mediator of consumer behavior. In general, I agree with this, however, we also need a publicly agreed upon conceptual definition of "involvement" and a valid measure of it. Then, in order to use "involvement" as a basis for a product classification system, it needs to be demonstrated that there is little within product variance across individuals on a valid measure of "involvement" and that different product classifications based on this measure imply different marketing programs.

#### Dependent Variable

In examining "involvement" as a mediator of consumer behavior, the author defines "consumer behavior" on a unidimensional scale anchored by a description of extensive problem solving and routinized response behavior. In responding to this scale, the respondent first reads a general description of extensive problem solving and routinized response behavior and then indicates for a number of different product purchase situations his or her hypothesized shopping behavior. This dependent measure is suspect for a number of reasons. First, I question whether shopping behavior can be validly measured on a unidimensional scale. When amount of stored information, usage of stored information, amount of information search from different sources and brand choice behavior are included as aspects of shopping behavior, the scale must be multidimensional. It is also not clear conceptually why this measure of "consumer behavior" was selected. All purchases of "low involvement" products should not necessarily exhibit routinized response behavior and all purchases of "high involvement" products should not necessarily exhibit extensive problem solving. Second, we should be suspicious of all self-report measures of shopping behavior. The study by Newman and Lockeman (1975), for instance, suggests that individuals generally give biased estimates of their information search behavior. This problem is compounded in this study since the measure is concerned with shopping behavior in a hypothetical as opposed to an actual situation. There is an indication that there are problems with this scale since according to Figure 1 in the paper, thirty to forty percent of the sample reported that they would use extensive problem solving to purchase lightbulbs, bread and toothpaste for their personal use.

#### First Analysis

In this analysis a repeated measure ANOVA is used to examine individual and product-in situation effects on the self-report measure of shopping behavior. Three of the product-in-situation are hypothesized to be "high involvement" situations. Although the classification of the pro-

duct-in-situation has face validity, it is somewhat surprising that the author did not actually test whether these product-in-situations actually differed on the two scales that the author uses to measure "involvement".

There is also some discussion in the paper about the difficulty in separating the effect of the interaction from the error in the analysis. It should be noted, however, that all of the remaining degrees of freedom in the analysis are accounted for by the interaction.

The results of the analysis indicates that approximately 17% of the variance in the dependent variable is accounted for by individual effects, approximately 20% by the products-in-situation factor and the remainder by the interactions. The author interprets this result as an indication that the respondents were "relatively homogeneous in behavior within a product class and heterogeneous across product". Based on the reported results this conclusion is difficult to justify. The results seem to indicate considerable heterogeneity within a product-in-situation. Unfortunately, it is difficult to ascertain how much of this is due to measurement problems and how much is due to actual behavioral differences.

#### Second Analysis

In this analysis, multiple regression analysis is used to examine the effect of "involvement" on the self-report shopping behavior scale after controlling for "knowledge about the product class", "amount of personal experience with the product class" and various demographic variables. In the analysis, "involvement" is hypothesized to consist of two components: (1) the "importance" of the purchase in attaining personal values and (2) the amount of "commitment" that the individual has to a particular brand within the product category. The validity of these components as a measure of "involvement" seem to be based on the results of a previous study (Lastovicka and Gardner, 1978). In this study the two dimensions from a multidimensional scaling of the relative importance of a number of different product categories were interpreted as "importance" and "commitment". The resulting configuration, however, indicates that the solution should probably have only a single dimension and the authors admit that this solution would be acceptable based on the resulting stress value. It should also be noted that according to the authors' interpretation of the two-dimensional solution, automobiles and houses would be lower on "commitment" than tissues. Consequently, there is some question as to whether or not these are valid measures of product class "involvement".

The results of the second analysis indicates that the set of independent variables explain approximately 16% of the variance in the dependent variable. The variable that is the most important predictor of the dependent variable, noncommitment, explains approximately 3% of the variance. This is especially disturbing because we would intuitively expect a large negative correlation between noncommitment and the dependent variable. An individual that is committed to a particular brand in a product category should exhibit routinized response behavior. This appears to be another indication that there are probably serious validity problems with the measures used in this study.

In summary, although I generally agree with the overall direction of this research, the author has not been as careful as he might have been in developing valid measures of his constructs. This appears to be especially true of the dependent variable. Consequently, the results of this study, which seems to indicate considerable heterogeneity of shopping behavior within product-in-situations and little relationship between "involvement" and shopping behavior, should not be considered conclusive. With better measures of "involvement" and valid multidimensional meas-

ures of shopping behavior, I suspect that we may find some rather strong relationships.

#### Future Research Directions

There seems to be general agreement on the potential importance of the construct "involvement" as a mediator of consumer behavior. However, there seems to be little agreement on precisely what "involvement" is or how it should be measured. In this section, I would like to present my conceptualization of "involvement" and then discuss future research directions using field and laboratory methods.

Involvement: What is it? Briefly, I view "involvement" as an individual level, internal state variable that indicates the amount of arousal, interest or drive evoked by a particular stimulus or situation. Involvement, therefore, has two dimensions, intensity and direction. Intensity concerns the level of arousal, interest or drive and direction concerns the evoking stimulus object and/or situation. In addition, the direction component may concern situations and stimulus objects at different levels of generality. Consequently, we may discuss "involvement" with respect to a product category, a particular brand and the purchase of a product for a particular reason.

For cognitive psychologists who are not used to thinking in terms of arousal or interest, "involvement" might be thought of in terms of a goal hierarchy (Simon, 1967). Here "involvement" might indicate the location of the particular stimulus and/or situation in a goal hierarchy at a particular point in time. A high "involvement" situation would have a reasonably high priority in the hierarchy and a low "involvement" situation a low priority. Consequently, we might think of "involvement" as mediating the amount of cognitive effort that will be expended in a given situation.

In the literature, two different types of definitions/measures of "involvement" have appeared. In one case, "involvement" has been viewed as a state variable and in others as a process. As discussed previously, I prefer to think of "involvement" as a state variable, so I will, therefore, discuss alternative definitions of this type first.

Most of the state variable measures of "involvement" that have appeared in the consumer behavior literature have their roots in social psychology. Here, starting with Sherif and Cantril (1947), it has been conceptualized in terms of the relationship between an individual's values and an issue or object. The more the issue or object becomes integrated with the individual's values, the higher the level of "involvement". This definition of "involvement", termed "ego involvement", has generally been viewed in the social psychological literature as a mediator of persuasion effects. More recent definitions of "ego involvement" have stressed "commitment" to the issue (Freidman, 1964) or the relationship of the issue or object to other elements of an individual's cognitive structure (Ostrom and Brock, 1968).

In developing measures of "involvement", researchers in consumer behavior have tended to simply apply social psychological measures of "ego involvement" with a particular issue to marketing situations involving products. In many cases, however, there appear to be conceptual problems in making this transition. For instance, Newman and Dolich and others (e.g. Lastovicka and Gardner, 1978), have adapted the Our Category Procedure (Sherif and Sherif, 1967), to measure "involvement" with a product category. Under this adaptation, the relative number of acceptable and unacceptable brands within a product category are used to measure "involvement" with the product category. Conceptually, however, this is not a measure of product class "involvement" as defined here. As mentioned previously, this

measure may be influenced by where the individual is in the purchase decision process if the product category is a consumer durable. If the product category is a consumer nondurable, an individual may have only one or two acceptable brands in order to reduce the amount of effort required in decision making in a "low involvement" product category. Consequently, the relative number of acceptable and unacceptable brands in product category is not a conceptual measure of product class "involvement." It may be a consequence of it, however, this is an empirical question.

Similarly, the "commitment" measure used by Lastovicka does not appear to be a conceptual measure of product class "involvement." Instead, it appears to be closer to a measure of the cognitive dimension of brand loyalty (Mitchell and Olson, 1975). Conceptually, I see little relationship between product class "involvement" and brand loyalty, however, again an empirical relationship may exist.

The second measure of "involvement" used by Lastovicka seems to be similar to "ego involvement" as defined by Sherif and Cantril (1947), however, it is different from my conceptualization given earlier. I view "ego involvement" as a possible antecedent condition of "involvement." As such it may be a sufficient condition for "high involvement," but it is not a necessary condition. For instance, a refrigerator may not be high on "ego involvement", however, for many people it may be a "high involvement" purchase situation.

Process definitions/measures of "involvement" usually involve information acquisition processes and decision processes. For instance, Krugman (1967), measures "involvement" in terms of the number of "connections" made between the product being advertised and the individual's personal life during exposure to an advertisement. Houston and Rothschild (1978) suggest a measure of "involvement" in terms of decision processes and Ray et al (1973) define "involvement" in terms of the orders or mental states after exposure to information. However, if we conceptualize "involvement" as discussed here, defining "involvement" as a process may be misleading. Although, the level of "involvement" may have a strong effect on the process used to make decisions or acquire information, I do not believe that it is the only determinant of these processes. For instance, in acquiring information from advertisements, the modality, the structure and content of the advertisement may also have an effect on these processes. Consequently, I don't believe it is useful to define "involvement" as a process.

Once we have arrived at a conceptual definition of involvement, there are a number of research directions that are important to pursue. These are best discussed under the classifications of field and laboratory studies.

Field Studies: The first priority here would be to develop a scale for measuring "involvement". Following the conceptual definition given here, a bank of questions might be developed to measure the level of arousal or interest in a particular product class or brand. "How likely are you to read a magazine article about \_\_\_\_\_?" or "Do you generally read advertisements for \_\_\_\_\_?" might be examples of these questions. Standard procedures, of course, should be used to insure the internal consistency, reliability, and the validity (Cronbach and Meehl, 1955; Cambell and Fiske, 1959) of the scale.

Once a reliable and valid scale is developed, there are a number of possible research directions. All of these involve the construction of a nomological net relating "involvement" to other aspects of consumer behavior (Cronbach and Meehl, 1955). First, the antecedent and consequent conditions of product class "involvement" should be examined. For instance, how exactly does "in-

volvement" affect purchase behavior? Second, what is the relationship between "involvement" and other constructs in consumer behavior such as perceived risk? Finally, how much between respondent variance is there on "involvement" within different product categories?

Laboratory Studies: The first priority here would be to examine how involvement affects information acquisition and decision processes. The central problem in these types of studies is how to create different "involvement" levels in the laboratory. This, I'm afraid, may prove to be a difficult task. In our current studies examining advertising effects at the individual level, we have, instead, tried to hypothesize how "involvement" may affect the type of processing that occurs during exposure to an advertisement (Mitchell, Russo and Gardner, 1978). We have hypothesized, for instance, that "involvement" may affect the strategy that an individual uses in processing information and the likelihood that competing stimuli will reduce attention levels. We have then attempted to create experimental conditions where these different processes will occur. This allows us to examine the effects of these different processes on the evaluations formed and the amount of information recalled about the brand and the advertisement.

#### SUMMARY

The concept of "involvement" seems to be potentially an important mediator of consumer behavior. However, before its potential can be determined empirically we need a publicly acceptable conceptual definition of "involvement," a valid scale for measuring it and methods for manipulating "involvement" in laboratory settings. In this paper, it has been defined as a individual level, state variable that measures the amount of arousal or interest in a stimulus object or situation. As such, "involvement" has two dimensions, intensity and direction. Consequently, we may talk about the amount of "involvement" with a product class, a brand or a purchase situation. Different measures of "involvement" that have appeared in the literature were then examined using this definition. Problems were indicated with each method.

Future research directions were also discussed. These involve the development of a scale for measuring "involvement," an examination of the antecedents and consequences of product class involvement and an examination of the effect of involvement on information acquisition and decision processes.

#### References

- Anderson, R.E. "Consumer Dissatisfaction: The Effect of Disconfirmed Expectancy on Perceived Product Performance," Journal of Marketing Research, 10(1973), 38-44.
- Campbell, D.T. and D.W. Fiske, "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix," Psychological Bulletin, 56(1959), 81-105.
- Cronbach, L.J. and P.E. Meehl, "Construct Validity in Psychological Tests," Psychological Bulletin, 52(1955), 281-302.
- Cronbach, L.J. and L. Furby, "How we should Measure 'Change'-or Should We?" Psychological Bulletin, Vol. 74 (1970), 68-80.
- 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: John Wiley & Sons, 1967, 257-266.
- Fishbein, M., Consumer Beliefs and Behavior with Respect to Cigarette Smoking: A Critical Analysis of the Public Literature, Staff Report of the Federal Trade Commission, May, 1977.
- Fishbein, M. and I. Ajzen, Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research, Reading, Mass.: Addison-Wesley Publishing Co., 1976.
- Freedman, J.L. "Involvement, Discrepancy, and Change," Journal of Abnormal and Social Psychology, 69(1964), 290-295.
- Houston, M.J. and M.L. Rothschild, "Conceptual and Methodological Perspectives on Involvement" in S. Jain, ed., Research Frontiers in Marketing: Dialogues and Directions, 1978 Educators' Proceedings, Chicago: American Marketing Association, 1978, 184-187.
- Krugman, H.E., "The Impact of Television Advertising: Learning Without Involvement," Public Opinion Quarterly, 29(1965), 349-346.
- Krugman, H.E., "The Measurement of Advertising Involvement," Public Opinion Quarterly, 30(1967), 583-596.
- Lastovicka, J.L. and D.M. Gardner, "Low Involvement Versus High Involvement Cognitive Structures," in H.K. Hunt, ed., Advances in Consumer Research, Vol. V, Ann Arbor: Association for Consumer Research, 1978, 87-92.
- Lastovicka, J.L. and D.M. Gardner, "Components of Involvement," in J.C. Maloney and B. Silverman, eds., Attitude Research Plays for High Stakes, Chicago: American Marketing Association (forthcoming).
- Lord, F.M., "Elementary Models for Measuring Change," in C.W. Harris, ed., Problems in Measuring Change, Madison" University of Wisconsin Press, 1963, 21-38.
- McLaughlin, B., "'Intentional' and 'Incidental' Learning in Human Subjects: The Role of Instructions to Learn and Motivation," Psychological Bulletin, 63(1965), 359-376.
- Mitchell, A.A. and J.C. Olson, "The Use of Restricted and Unrestricted Maximum Likelihood Factor Analysis to Examine Alternative Measures of Brand Loyalty" in E.M. Mazze, ed., Marketing: The Challenges and Opportunities, Chicago: American Marketing Association, 1975, 181-186.
- Mitchell, A.A., J.E. Russo and M. Gardner, "Strategy Limitation Effects on the Learning of Advertising Messages," Working Paper 49-77-78, Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh, Pa. 15213.
- Newman, J.W. and B.D. Lockeman, "Measuring Prepurchase Information Seeking," Journal of Consumer Research, 2 (1975), 216-222.
- Olson, J.C., "The Stability of Responses Obtained by Free Elicitation: Implications for Measuring Attribute Salience and Memory Structure," in William Wilkie, ed., Advances in Consumer Research, Vol. VI, (forthcoming).
- Olson, J.C. and P.A. Dover "Disconfirmation of Consumer Expectations Through Product Trial," Journal of Applied Psychology (forthcoming).
- Ostrom, T.M. and T.C. Brock, "A Cognitive Model of Attitudinal Involvement," in R.P. Abelson, et al, eds., Theories of Cognitive Consistency: A Sourcebook, Chicago: Rand McNally, 1968.
- Postman, L., "Short-term Memory and Incidental Learning,"

in W.A. Melton, ed., Categories of Human Learning, New York: Academic Press, 1964.

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, Ca.: Sage Publications, 1973.

Robertson, T.S., "Low Commitment Consumer Behavior," Journal of Advertising Research, 16(1976), 19-27.

Rothschild, M.L., "Political Advertising: A Neglected Policy Issue in Marketing," Journal of Marketing Research, 15(1978), 58-71.

Rothschild, M.L. and M.J. Houston, "The Consumer Involvement Matrix: Some Preliminary Findings," in B.A. Greenberg and D.N. Bellenger, Contemporary Marketing Thought, 1977 Educators' Proceedings, 1977, 95-98.

Sherif, M. and H. Cantril, The Psychology of Ego Involvement, New York: John Wiley & Sons, 1947.

Sherif, M. and C.J. Hovland, Social Judgment: Assimilation and Contrast Effects in Communication and Attitude Changes, New Haven: Yale University Press, 1953.

Sherif, M. and C.W. Sherif, "The Own Categories Procedure in Attitude Research," in M. Fishbein, ed., Readings in Attitude Theory and Measurement, New York: John Wiley and Sons, 1967.

Sherif, C.W., M. Sherif and Nebergall, Attitude and Attitude Change: The Social Judgment Involvement Approach, New Haven" Yale University Press, 1965.

Simon, H.A., "Motivational and Emotional Controls of Cognition," Psychological Review, 74(1967), 29-39.

Michael L. Ray, Stanford University

## Abstract

The papers in this session reflect an advanced stage of use of involvement in consumer research. Serious questions about reliability and particularly validity and usefulness of the concept are being asked. The answers seem to be that involvement measures can be reliable, valid and useful only within certain situations. At present, as the title indicates, involvement has been shown to be measured well and to be valuable only in the communication effects situation. And even there it must be used along with other variables. Some general implications for the development and use of variables such as involvement are discussed.

## Introduction

The modern introduction of ego involvement as a social psychological variable probably can be traced to the book by Sherif and Cantril (1947) published over three decades ago. Although the concept appeared sporadically in the consumer research literature since then, it was introduced most effectively to our field only about thirteen years ago by Krugman (1965) when he explained the effects of television advertising with a low involvement viewer hypothesis. Since that time the concept has been expanded for the consumer field by Ray et al (1973, Ray 1976), Robertson (1976), Rothschild (1974), and by a series of authors who participated in an American Marketing Association Attitude Research conference devoted primarily to the involvement variable (Maloney and Silverman, 1978).

Given its brief history in consumer research, the involvement concept has only just started to be questioned and refined for its application to consumer behavior issues. The papers in this session reflect this stage in the application of the concept.

## The Present and Potential State of Involvement in Consumer Research

The main questions now seem to be ones of validity. What is the best measure of involvement? What are the components of involvement? Does involvement affect consumer behavior in the way involvement "theorists" say it should? What are the alternative concepts and measures of concepts which might operate to explain consumer behavior phenomena in conjunction with or in place of involvement?

It is a basic theme of this discussion paper that the ultimate answers to these questions about involvement will be quite conservative and limiting. That is, there will be no one best method to measure involvement. The components and effects of involvement will be governed primarily by the consumer decision situation. And involvement alone will not explain consumer phenomena. Other mediating variables will have to be invoked. In short, involvement will not be the whole answer to the fascinating problems of consumer research. Involvement will provide part of the answer only in certain situations. At this point, as the title of this paper indicates, we can be somewhat confident in the operation of involvement along with other variables only in the communication situation.

This may seem to be a pessimistic view. Actually it is realistic. Consumer research tends to be an applied field. We borrow concepts and techniques

from several sciences. Early in the process we have unrealistic expectations for a new concept. Then, because these expectations cannot possibly be met, there is a period of disillusionment. That period could lead to the concept being dropped or being refined. Fortunately, as illustrated by the papers in this session, the period of disillusionment for the involvement concept is leading to refinement rather than rejection.

## The Papers

## Lastovicka

Lastovicka asks all the right questions about involvement and provides evidence of its nature, importance and interaction with other variables. It is clear that involvement as conceptualized by Lastovicka is an important variable. It is also just as clear that involvement alone will not explain consumer behavior; other variables must be used.

There are limits to the inferences we can make from the Lastovicka research, however. These limits are caused basically because his key measures have not been tested for reliability or validity. While Lastovicka is in the majority of consumer research in this regard (e.g. Ray, 1979) and has actually done some previous validity testing of measures related to involvement (Lastovicka and Gardner, 1978), the study is based entirely on responses on a five-point scale to two previously untested descriptions. That yeasaying is present is clear from the mean results which show that even the low involvement product situations produced scores around the mid-point of the scale. It appears that the respondents as a group didn't really want to admit that their behavior was totally to the routinized behavior side of the continuum.

This lack of validation for the key measure could be dismissed if this were only the first of a series of studies. But in consumer research we do not have a tradition of replications.

It could also be argued that the possible yeasaying or role-playing decreased variance so that the strong effect of involvement shown in the study was actually conservative. But this argument ignores the fact that, as mentioned above, the two descriptions consumers rated for seven product categories were untested and, I believe, biased toward some effect of involvement. My belief should have less credibility than Lastovicka's here since this is his study. But on the other hand anybody's guess is still a guess since his key measure was not validated.

What is clear is that there is large residual error variance which cannot be explained by either subject or product (involvement). My guess is that an error term of this size is probably due to the interaction of the products with the yeasaying or role playing biases that probably exist in this study. This is not to say that Lastovicka's results are worthless. But they must be followed with much further work if they are to have lasting value.

Some hint as to the character of this further work is seen in the limits to Lastovicka's "Implications"

section. While I agree that his research can support his statement that "... no one classification device is a very satisfactory device for totally accounting for differences in acquisition behavior," I find his speculations about the information processing behaviors of people new to a product class to be completely without foundation in his research and counter to results my coworkers and I have obtained. Specifically, I find it hard to believe that there are substantial numbers of people purchasing in any product category which, as Lastovicka implies, never do any extensive problem solving. I guess I would be classified as a proponent of "low involvement theory" as Lastovicka put it. Even though I believe this is a misuse of the word "theory," I do believe that we "theorists" never said that a consumer's whole experience with a product would be in a low involvement mode.

In sum, Lastovicka's study is a nice demonstration of the power of involvement and a first step which points to the need for further research. That further research should consist of experimental studies which determine how people in various stages of the decision process behave with regard to various sources of information. Notice it will be necessary to determine how people actually behave rather depending on how they say they behave. Thus far such studies have been done only with regard to advertising, with the exception, perhaps, of the next study to be discussed.

Newman and Dolich

This is an interesting application of Sherif's ego-involvement notions to consumer behavior; particularly in light of the fact that this was a high-involvement communication (a live product demonstration of an automobile) and that the study was experimental in nature. The author's conclusion that the operation of the ego-involvement was as expected (if it can be assumed that the demonstration was a negative experience) is both provocative and full of implications for marketers.

Unfortunately there are several shortcomings to the Newman and Dolich research which limits the implications we can draw from it. Like Lastovicka, they use scales that have not been tested for reliability and validity for this application. The sample sizes are extremely small, apparently ranging from about five to 12 in the six cells of the design represented in Table 1. Since respondents were arbitrarily divided into two groups on the basis of their pretest scores, there is the slim possibility of artifactual results caused by "regression toward the mean" (Campbell and Stanley, 1966). The curves in Figure 2 which we are to compare with the Sherif predictions of Figure 1 are not truly curves at all since they are based on only two data points. Finally, and most importantly, there is no statistical analysis so we have no way of knowing whether the differences reported in the paper are true differences in a statistical sense.

None of these deficiencies in the paper are serious in themselves. But, taken together, they weaken my faith in Newman and Dolich's conclusions. This is unfortunate because unlike recent papers in this area (e.g. Lastovicka and Gardner, 1978, Rothschild and Houston, 1977) this study is experimental and observes actual response rather than just consumer opinion about their response. Also it is one of the first experimental studies of this sort using a high involvement communication (another example: Swinyard and Coney, 1978). What can be said, however, is that the results do fit theory, and the method should be used again with the corrections implied above.

Grønhaug and Kangun

This paper is both more and less than a study on involvement. It is "more" in the sense that it offers non-American data on a socially-important issue in consumer behavior. It is "less" in the sense that we have to assume that smokers and non-smokers and those who have tried to stop actually differ in involvement. Grønhaug and Kangun would not, I believe, presume that this difference truly exists. They would probably agree that smokers and non-smokers differ on many more variables than involvement. The authors cannot be blamed for the placement of their paper in a session on involvement. Nor can I be blamed for having difficulty in discussing it in this regard. Perhaps our program chairman, William Wilkie, is the person we should blame. I certainly wouldn't do this, however.

At the opportunity of being repetitive, let me say that this paper, like the previous two, suffers because measures are of unknown reliability and validity. For example, the results on "awareness" which are given much importance by the authors, are based on a recognition measure. It has long been established (e.g. Wells, 1964) that recognition measures lead respondents to give subjective estimates of the probability that they were exposed to some communication rather than their true "awareness" of it. So when Grønhaug and Kangun find, to their surprise, that smokers and non-smokers are equally aware of statements on smoking and health, they are simply replicating the biased sort of result that has been found with recognition measures through the years.

Finally, in either reanalyses of the present data or in future studies in this area, Grønhaug and Kangun should consider the sort of involvement measure proposed by Rothschild and Houston (1977). The results would probably indicate both high and low involvement segments.

#### Implications

Involvement was introduced to consumer research by Krugman to explain the effects of television advertising. Although he presented no direct data to support his hypothesis, subsequent experimental studies (e.g. Ray, et al., 1973, Ray, 1976, Ray and Webb, 1976, Rothschild, 1974, Swinyard and Coney, 1978) indicate that advertising and particularly television advertising operates the way Krugman said it does, partially because of involvement but also because of a variety of other factors including the nature of the medium, the perceived degree of brand differentiation, the sources of information most important in decision-making, and the degree of experience consumers have had with the product category.

Nothing in the present set of papers or any others I've seen recently on involvement in consumer research convinces me that a single measure of involvement will be found for all consumer research situations, that involvement alone can be used to explain any consumer behavior phenomenon or that even in that limited role it can do anything more than explain communication-advertising effects.

Measures and applications of involvement should be developed in individual consumer research application situations. An excellent example of such development of the concept in the political campaigning area was done by Rothschild (1974, Rothschild and Ray, 1974). He based his work on Krugman's ideas, political voting studies and the results of a large scale multiple measures study of political attitudes and behavior (see a report in Heeler and Ray, 1972). Recently his conceptualization was upheld in a study by Swinyard and



Coney (1978). But Rothschild and Houston's attempt to find a more general measure of involvement produced less clear-cut results.

Another exemplary use of involvement was by Webb (1978, Ray and Webb, 1976) in the area of television clutter. Using Krugman's "connections" measure of involvement he selected commercials that were high and low in involvement. He then used these commercials in an experimental study with variations of television clutter. He found that the high involvement commercials (which probably scored high on the connections measure for a variety of reasons other than "pure" involvement) were much less affected by the clutter conditions than were the low involvement commercials. While this study may not contribute a great deal to the "theory" of involvement, it does provide valuable information to those who are attempting to develop advertising for a difficult television environment. And here Webb used a measure of involvement which, while imprecise, was developed by Krugman for use with advertising and which could gauge all the important aspects of involvement for this situation of interest. And I believe that Webb's experiment provides some reliability (in the prestudy) and validity information on the measure and concept precisely because the identified advertisements operated the way Krugman would say they would.

This discussion reflects trends that are obvious in consumer research and should be applied in the development and use of the involvement concept. First, consumer behavior is highly situational. Involvement up to now has been shown to work in one broad type of consumer situation: information processing of communications, particularly advertising. Successful uses of the concept have been careful long-term efforts in particular communication-advertising situations.

Second, the causes of consumer behavior are multivariate. Involvement has a number of components depending on the situation, and involvement alone will not explain consumer behavior. Other mediating variables are necessary.

Third, reliability and validity testing in consumer research is inadequate, although it is improving. The papers in this session, although sometimes weak in measures, provided some experimental evidence of the validity of the involvement concept and its measures.

Fourth, it is better to observe consumer responses in experiments than to ask consumers to guess how they would respond. And the experimental method is a good one for obtaining validity indications for concepts to be used in consumer research.

#### References

- Donald T. Campbell and Julian Stanley, Experimental and Quasi-Experimental Designs for Research (Chicago: Rand-McNally, 1966).
- Roger M. Heeler and Michael L. Ray, "Measure Validation in Marketing," Journal of Marketing Research, 9(November 1972), 361-71.
- Herbert E. Krugman, "The Impact of Television Advertising: Learning without Involvement," Public Opinion Quarterly, 29(Fall 1965), 349-56.
- John L. Lastovicka and David M. Gardner, "Components of Involvement," In John C. Maloney and Bernard Silverman (Eds.) Attitude Research Plays for High Stakes (Chicago: American Marketing Association, 1978).
- John C. Maloney and Bernard Silverman (Eds.) Attitude Research Plays for High Stakes (Chicago: American Mar-

keting Association, 1978).

Michael L. Ray, "Attitude as Communication Response," In Deborah Johnson and William D. Wells, Attitude Research at Bay (Chicago: American Marketing Association, 1976).

Michael L. Ray, "Measurement and Marketing: Is the Flirtation Going to Lead to Romance?" Journal of Marketing Research, 16(February 1979).

Michael L. Ray, Alan G. Sawyer, Michael L. Rothschild, Roger M. Heeler, Edward C. Strong, and Jerome B. Reed, "Marketing Communication and the Hierarchy of Effects," In Peter Clarke (Ed.) New Models for Mass Communication Research (Beverly Hills, Ca.: Sage Publishing, 1973), 147-76.

Michael L. Ray and Peter H. Webb, Experimental Research on the Effects of Television Clutter: Dealing with a Difficult Media Environment (Cambridge, Mass.: Marketing Science Institute, 1976), Report No. 76-102.

Thomas S. Robertson, "Low Commitment Consumer Behavior," Journal of Advertising Research, 16(April 1976), 19-27.

Michael L. Rothschild, "The Effects of Political Advertising on the Voting Behavior of a Low-Involvement Electorate," unpublished Ph.D. dissertation, Graduate School of Business, Stanford University, 1974.

Michael L. Rothschild and Michael J. Houston, "The Consumer Involvement Matrix: Some Preliminary Findings," In Barnett A. Greenberg and Danny N. Bellenger (Eds.) Contemporary Marketing Thought (Chicago: American Marketing Association, 1977), 95-8.

Michael L. Rothschild and Michael L. Ray, "Involvement and Political Advertising Effect: An Exploratory Experiment," Communication Research, 1(July 1974), 264-84.

Muzafer Sherif and Hadley Cantril, The Psychology of Ego Involvements (New York: Wiley, 1947).

William R. Swinyard and Kenneth A. Coney, "Promotional Effects on a High- versus Low-Involvement Electorate," Journal of Consumer Research, 5(June 1978), 41-8.

Peter H. Webb, "Consumer Information Acquisition in a Difficult Media Environment: The Case of Television Clutter," Unpublished Ph.D. dissertation, Graduate School of Business, Stanford University, 1978.

William D. Wells, "Recognition, Recall, and Rating Scales," Journal of Advertising Research, 4(September 1964), 2-8.

P. S. Raju, The Pennsylvania State University

Abstract

This paper focuses on stimulus variables, response variables, and interrelationships between them in relation to new product research. Results from a study correlating certain specific stimulus variables and response variables are presented. Four stimulus variables, one cognitive response variable, and four behavioral response variables were included in the study. Implications of the study are discussed and issues for future research are identified.

Introduction

New product research is an important area for most companies today and a substantial amount of the marketing literature has been devoted to this area. Often quoted statistics with regard to new product failures serve to emphasize the importance of understanding consumers' responses to new products. Most available evidence in the area evolves from the "diffusion of innovations" tradition. However, on reviewing the literature in this area it is evident that indepth studies of consumers' responses to new products and factors influencing these responses have rarely been conducted. Recently, research in the area of "variety seeking" has offered further insights into factors that influence response to new products (Venkatesan, 1972; Raju, 1977). Theories in this area have identified several useful variables. Unfortunately rigorous empirical testing of these theories has proved difficult thus far.

The need to systematically examine factors that influence response to new products cannot be denied. The present paper is a first step in that direction. More specifically, the present paper has the following objectives: 1) to offer insights on stimulus variables, response variables, and their interrelationships in new product research, and 2) to identify issues that could be addressed by future studies in the area.

Stimulus-Response Variables

The stimulus-response (S-R) model offers the simplest way of conceptualizing consumer reaction to new products. In this model response is a function of stimulus characteristics. Individual characteristics will be incorporated only in so far as the stimulus variables are consumer perceptions of the stimulus rather than absolute measures. Specific individual characteristics (such as personality, life style, etc.) and other factors that moderate the stimulus-response relationship will not be explicitly considered in this paper. Interest, therefore, will center primarily on stimulus variables, response variables, and their interrelationships.

New products can be characterized in terms of several stimulus variables. The "diffusion of innovations" literature has identified factors which influence responses to new products such as relative advantage, compatability complexity, divisibility, and communicability (Robertson, 1972). Other variables such as novelty, ambiguity, incongruity, and surprisingness are recognized as being important in the "variety seeking" literature (Berlyne, 1960; Raju, 1977). In addition, marketing researchers have identified some stimulus variables, the most noteworthy of which is perceived risk (Bauer, 1960; Roselius, 1971).

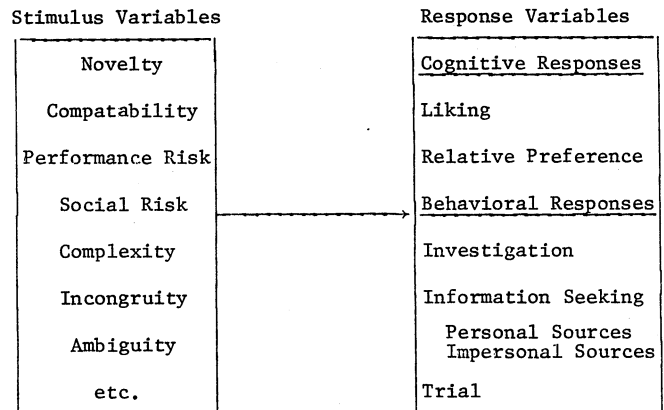
Consumer's responses to new products can be manifested in different ways. An exhaustive classification is not

possible but one could classify responses basically into two categories: "cognitive" and "behavioral." In this paper cognitive response is considered to be an overall emotional reaction toward the new product and is operationalized as relative preference for the new product over existing alternatives. Three types of behavioral response are considered: investigatory response, information seeking response, and trial response. Investigatory response pertains to exposing oneself more to a new product by looking, moving toward, or examining the new stimulus. Such responses which are "exploratory" in nature have been incorporated in the "variety-seeking" literature (Berlyne, 1960; Raju, 1977). Information seeking response comprises of acts performed to obtain more information from personal or impersonal sources. The former relates to word of mouth communication with friends, relatives, salespeople, etc., whereas the latter relates to information seeking from the product package, advertisements, and mass media. Finally, the trial response incorporates purchase of the new product. The three behavioral responses identified could be performed by a consumer to first investigate a product, then seek information about it, and ultimately try it if she is pleased with the results of her previous responses. However, a strict adherence to such a hierarchy in all new product situations would not be expected.

The stimulus variables and response variables in a new product situation are summarized in Figure A. A study that was designed to address some probing questions regarding interrelationships between these variables will now be discussed. Other important issues in the area will also be briefly reviewed later in this paper.

FIGURE A

Stimulus and Response Variables



Study Description

Objectives

The study was designed to examine a few specific stimulus and responses variables in new product situations and interrelationships between these variables. The domain of consumer packaged goods and small appliances was chosen for the study. Four stimulus variables were included, namely, novelty, incompatability, performance risk, and

TABLE 1

## Operationalization of Stimulus Variables

social risk. Response variables included were relative preference, investigatory response, information seeking from personal sources, information seeking from impersonal sources, and trial. Specific questions addressed in the study were: 1) Are the stimulus variables correlated with each other or do they represent independent dimensions? 2) Are the response variables correlated with each other or do they represent independent dimensions? 3) Are particular stimulus variables correlated with particular response variables? 4) Do certain stimulus variables play a major role in determining response to new products?

Only aggregate relationships in the form of correlations have been emphasized in this study. In that respect it should be considered an exploratory study since no attempt has been made to test specific forms of relationships among the variables. However, some helpful conceptual ideas for this purpose have been offered on the basis of the results from this study.

## Sample

Respondents in the study were forty women selected on the basis of a convenience sample, whose cooperation was obtained through womens' clubs in the area. Monetary incentive was provided to the clubs to encourage participation by their members. Respondents ranged from 22 to 67 years of age with the median age being 34 years. They were almost equally divided between the categories of full time employment, part-time employment, and full time homemakers. Finally, ninety percent had graduated from high school, with about thirty percent having a college degree. The sample is thus somewhat biased toward the more educated consumer but well balanced in terms of age and employment.

## Method

Data on stimulus variables and response variables for twenty new products were collected by means of a questionnaire. The twenty new products were carefully selected. Several past issues of "Advertising Age" were monitored for announcements regarding new product introductions in the market and brief descriptions of these new products were also obtained. The new product alternatives generated were then screened and twenty products selected for inclusion in the study. The criteria used in screening the alternatives were, 1) the products should be of interest to most women, 2) a respondent should be able to decide by herself fairly quickly how she is likely to respond to the new product (products for which joint decision making was likely were therefore eliminated), 3) they should be consistent in terms of the general category to which they belong so that wide variations in price and other extraneous factors could be avoided, and 4) they should either not have been introduced in the market or should have been in the market only for a short period at the time of the study. The twenty new products selected on the basis of these criteria are listed with their descriptions in the Appendix.

Measurement of Stimulus Variables. A brief description of each new product was provided to respondents in the questionnaire. The descriptions were designed to provide just sufficient information to enable respondents to judge what the product would be like and how it was intended to function. After each new product description, measures of four stimulus variables were obtained. The four stimulus variables were novelty, incompatibility, performance risk, and social risk and these were operationalized in terms of respondents' perceptions of the new products. Other stimulus variables were considered but eliminated either because they were not relevant to the products considered or did not lend themselves to operationalization with product descriptions. Operationalization of the four stimulus variables is summarized in Table 1.

1. Novelty: How different is this product from other currently available products that you might use for the same purpose?  

not at all	extremely
different _____	different
2. Compatibility: Assume you have bought this product. How well would it fit in with your present life style or way of doing things?  

will fit in	will not at
very well with	all fit in
my life	with my
style _____	life style
3. Performance Risk: Sometimes buying products could be risky because they perform poorly or may not be of satisfactory quality. How much of this "performance or quality" risk do you think there is in buying a product of the kind described above?  

no perform-	extreme pre-
ance risk _____	formance risk
4. Social Risk: Sometimes buying products can be risky in a social sense, meaning that possession of them could be socially unacceptable because the product may not fit in with your life style, or friends and relatives may not approve of your buying the product. How much of this kind of social risk is there in buying a product of the kind described above?  

no social	extreme
risk _____	social risk

Measurement of Response Variables. One cognitive response variable and four behavioral response variables were also measured for each product description. The cognitive response variable was relative preference and the behavioral response variables were investigatory response, information seeking from personal sources, information seeking from impersonal sources, and trial. The cognitive response variable was measured after each product description. For measuring the behavioral responses, each response was described in turn and the likelihood of performing that response for each of the twenty new products was obtained. Operationalization of the response variables is shown in Table 2.

## Analyses and Results

Pearson product moment correlations were obtained between relevant variables for each product separately. Since perceptions and responses were of interest it was considered desirable to include only respondents who were in the potential market for each product in the analysis for that product. The information on whether each respondent could or could not use each of the twenty new products was obtained in the questionnaire for this purpose. Tables 3 to 6 show the correlations between different sets of variables. In each case the table provides the number and percentage of products that showed significant correlations at the 0.10 probability level as well as the range of these significant correlations. Each of these tables is briefly discussed below.

Table 3 shows correlations between the four stimulus variables. There were no consistently strong correlations between particular stimulus variables. A significant correlation between incompatibility and perceived risk (both performance and social) was shown by a fair number

of products. The lack of strong correlations between the stimulus variables is, perhaps, indicative of the fact that these variables represent independent dimensions of a stimulus. Thus, if a product rates high on one dimension such as novelty it need not also rate high on another dimension such as incompatibility.

TABLE 2

Operationalization of Response Variables

1. Relative Preference: Overall, how much do you prefer this product in comparison with currently available products that could be used instead?

much prefer existing alternatives \_\_\_\_\_ much prefer new alternatives \_\_\_\_\_

2. Investigatory Response: While doing your regular shopping, you see this product on the store shelf. How likely is it that you will pick the product up and investigate it further out of curiosity?

not at all likely \_\_\_\_\_ very likely \_\_\_\_\_

3. Information Seeking from Personal Sources: While visiting a good friend, you notice this product in her house. How likely is it that you will try to find out more about the product from your friend?

not at all likely \_\_\_\_\_ very likely \_\_\_\_\_

4. Information Seeking from Impersonal Sources: While reading a magazine at home, you notice an advertisement which announces the introduction of this product into the market. The ad has a removable section in it which you can send to the company to get more information about the product and a discount coupon for 10% off of the regular price. How likely is it that you will send for the additional information and the discount coupon for this product?

not at all likely \_\_\_\_\_ very likely \_\_\_\_\_

5. Trial Response: You notice a large display in the store announcing the introduction of this product and showing you the benefits of the product. The store's policy is that it will replace an item that is defective. How likely is it that you will buy the product?

not at all likely \_\_\_\_\_ very likely \_\_\_\_\_

Table 4 shows correlations of the cognitive response variable with stimulus variables and behavioral response variables. Among the stimulus variables incompatibility and novelty showed stronger relationships with relative preference, the correlations being significant for 40% and 35% of the products respectively. Incompatibility was correlated negatively with relative preference as would be expected. However, novelty was correlated positively with relative preference indicating that for the type of products considered, consumers prefer those that are more different from existing alternatives. Perceived risk, and especially performance risk was correlated with relative preference for few products. This may be due to the nature of products included in the study which were all consumer packages goods and small appliances. Perceived risk probably determines preference at the brand level for these products rather than at the product level. With respect to response variables, relative preference was more strongly related with trial and investigatory response. Sixty-five percent of the products showed a significant correlation between relative preference and

trial and forty percent showed a significant correlation between relative preference and investigatory response. Information seeking from personal sources and from impersonal sources showed significant correlations with relative preference only for 15% of the products.

TABLE 3

Correlations Between Stimulus Variables\*

	Novelty	Incompat-ability	Perf. Risk	Soc. Risk
Novelty	1.0			
Incompat-ability	4 pdts (20%) (0.444 to 0.507)	1.0		
Perf. Risk	2 pdts (10%) (0.426 to 0.522)	8 pdts (40%) (0.367 to 0.578)	1.0	
Soc. Risk	0 pdts	6 pdts (30%) (0.364 to 0.807)	4 pdts (20%) (0.414 to 0.834)	1.0

\*The table shows the number and percentage of products which showed significant correlations for each pair of variables at the 0.10 significance level. The range of these significant correlations is also shown in each case. Other correlation tables are interpreted in a similar manner.

TABLE 4

Correlations of Relative Preference with Stimulus Variables and Behavior Response Variables

	Novelty	Incompat-ability	Perf. Risk	Soc. Risk
Relative Preference	7 pdts. (35%) (0.386 to 0.578)	8 pdts. (40%) (-0.328 to -0.5952)	1 pdt. (5%) (-0.4412)	3 pdts. (15%) (-0.394 to -0.463)
	Inv. Response	Info. Seek- ing Personal Sources	Info. Seek- ing Impersonal Sources	Trial
Relative Preference	8 pdts. (40%) (0.3493 to 0.6761)	3 pdts. (15%) (0.4072 to 0.5953)	3 pdts. (15%) (0.5593 to 0.9300)	13 pdts. (65%) (0.3579 to 0.7316)

Correlations between response variables are shown in Table 5. These variables were significantly correlated between themselves for a fairly large number of products. Except for the correlation between information seeking from personal sources and from impersonal sources which was significant only for 35% of the products, all other pairs of variables showed significant correlations for at least 65% of the products. These results suggest that different behavioral responses may be dependent on each other in some systematic manner and that the performance of one response could lead to other responses. However, the weaker relationship between the two types of information seeking responses suggests that each may be relatively independent of the other even though both are related to other responses. This could happen, for example, if consumers prefer to seek information either from personal sources or from impersonal sources but rarely do both.

TABLE 5

## Correlations Between Response Variables

	Inv. Response	Info. Seeking Personal Sources	Info. Seeking Impersonal Sources	Trial
Inv. Response	1.0			
Info. Seeking Personal Sources	16 pdts. (80%) (0.398 to 0.973)	1.0		
Info. Seeking Impersonal Sources	14 pdts. (70%) (0.446 to 0.705)	7 pdts. (35%) (0.372 to 0.645)	1.0	
Trial	14 pdts. (70%) (0.350 to 0.816)	13 pdts. (65%) (0.361 to 0.836)	13 pdts. (65%) (0.404 to 0.756)	1.0

Table 6 shows correlations between stimulus variables and behavioral response variables. Considering each behavioral response variable separately it can be seen that the stimulus variable most related to it is incompatibility. In the case of the trial response, which is of most interest to marketers, incompatibility and novelty were the two most important variables, with perceived risk showing insignificant correlations for practically all products. It is surprising that although performance risk were significantly correlated with other response variables for a few products, those correlations were all negative. One would generally expect investigatory and information seeking responses to increase with perceived risk. A possible explanation for this is offered later.

TABLE 6

## Correlations Between Stimulus Variables and Behavioral Response Variables

	Inv. Response	Info. Seeking Personal Sources	Info. Seeking Impersonal Sources	Trial
Novelty	2 pdts. (10%) (0.3433 to 0.3625)	2 pdts. (10%) (0.5107 to 0.5591)	1 pdt. (5%) (0.4564)	4 pdts. (20%) (0.4083 to 0.6491)
Incompatibility	9 pdts. (45%) (-0.3585 to -0.7994)	7 pdts. (35%) (-0.4097 to -0.6568)	5 pdts. (25%) (-0.3137 to -0.6757)	10 pdts. (50%) (-0.3354 to -0.7894)
Perf. Risk	4 pdts. (20%) (-0.4698 to -0.9417)	3 pdts. (15%) (-0.4214 to -0.5554)	5 pdts. (25%) (-0.3131 to -0.4952)	0 pdts.
Soc. Risk	6 pdts. (30%) (-0.4079 to -0.8014)	2 pdts. (10%) (-0.5334 to -0.5505)	1 pdt. (5%) (-0.3108)	1 pdt. (5%) (-0.3499)

A summary of the results is as follows:

1. The stimulus variables are unrelated for most of the products considered and seem to represent independent dimensions.

2. Among the stimulus variables, incompatibility and novelty are most related to relative preference. Incompatibility is negatively correlated with relative preference and novelty is positively correlated.

3. Among the behavioral response variables, investigatory response and trial are most related to relative preference. Relative preference and trial show a strong relationship with significant correlations for 65% of the products considered.

4. Among the stimulus variables, incompatibility is most correlated with each of the behavioral response variables. Novelty also seems to have some influence on trial. Performance risk and social risk have no influence on trial but show negative correlations with other response variables for a few products.

5. The behavioral response variables show significant correlations between themselves for several products. However, information seeking from personal sources and from impersonal sources seem to be relatively independent although each is related to other behavioral response variables.

The next section will be devoted to analyzing the above results in terms of implications for marketing researchers.

## Implications and Conclusion

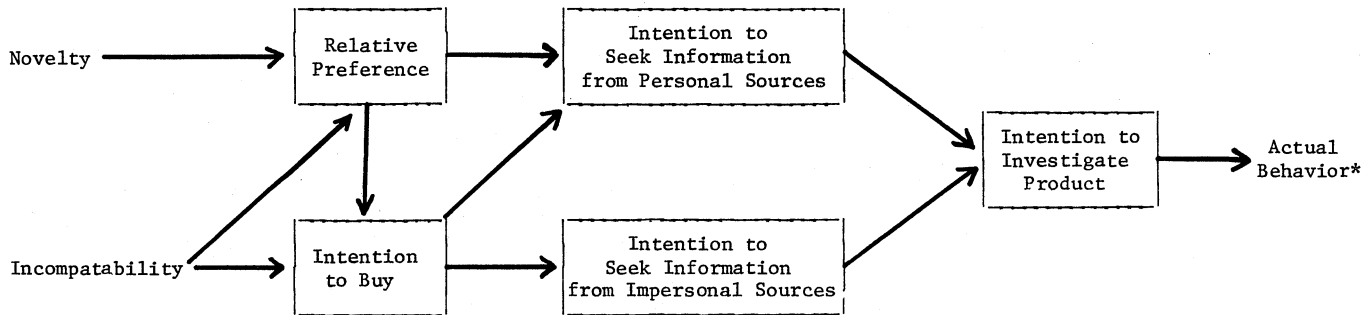
One major issue for further investigation is whether stimulus variables identified in the literature represent independent dimensions as suggested by this study for some of the variables. Future studies should include more stimulus variables such as complexity, incongruity, perceived conflict, etc. One problem with these variables is the difficulty in operationalizing them. Research effort should therefore be directed to operationalization problems also. In essence, the area of new product perception requires more research attention.

The correlations of performance risk and social risk with relative preference and behavioral response variables are interesting. Firstly, for products included in this study, perceived risk has very little influence on relative preference and behavioral response variables. This seems to point to the fact that risk by itself is not of much concern to consumers in these product categories. There are probably other important variables that determine their responses to new products. Secondly, for a few products perceived risk is negatively correlated with investigatory and information seeking responses. This seems to indicate that in cases where risk is important, consumers are passive with respect to this variable. The tendency of consumers is, therefore, to ignore the product rather than investigate it or seek information about it. If this is true it has important implications for marketing products, because the burden of reducing perceived risk falls completely on the marketer. The conclusion, however, may not be valid for products where all alternatives have about the same amount of risk or for product categories that consumers have great need for. Thus, in buying a car for the first time a consumer would be expected to investigate alternatives and actively seek information about them.

The stimulus variables incompatibility and novelty are also interesting in several ways. Since these variables are to some degree positively correlated (Table 3) but each is related in the opposite manner to relative preference, the implication for the marketer is that a new product should be designed to be optimally different from existing alternatives to obtain positive responses from consumers. The nature of the relationships between novelty, incompatibility, relative preference, and trial are particularly interesting. Both incompatibility and novelty are related to relative preference which in turn is correlated with trial. Thus, these stimulus variables

FIGURE B

A Causal Model of Stimulus-Response Relationships



\*In actual behavior the causal order shown for intentions may be reversed with investigation leading to information seeking which then may lead to actual trial of the product.

probably influence behavioral response by means of the cognitive response. In the case of novelty this notion is reinforced because its relationship with relative preference (Table 4) is somewhat stronger than its relationship with trial (Table 6). However, in the case of incompatibility its direct relationship with trial is somewhat stronger than its relationship with relative preference. This raises the possibility that incompatibility might influence willingness to try a product directly without the intermediate stage of relative preference. In generalizing this idea to several stimulus variables one could say that stimulus variables can influence willingness to try a product in two ways, either directly or through intermediate cognitive responses. The two ways are not mutually exclusive so that a particular variable may sometimes operate both ways.

Considering the relationship between cognitive response and behavioral response it can be seen from Table 4 that relative preference has the strongest relationship with trial. Table 5 shows that several of the response variables are also correlated among themselves for a fairly large number of products. This suggests the hypothesis that relative preference first influences the interest in trying a product which then influences other types of behavioral responses. This hypothesis is worthy of further consideration because it runs contrary to the general notion that a consumer, if he prefers a new product, will first investigate it, then seek information about it, and finally decide to try it.

The implications discussed above indicate that there are several issues that remain to be investigated in the area of new products. At the heart of these issues is the question of what causal linkages exist between stimulus variables, cognitive response variables, and behavioral response variables. One possible model of causal linkages between variables is shown in Figure B. The model shows the influence of two major stimulus variables, incompatibility and novelty, and treats them as independent dimensions. Both these stimulus variables influence intention to try the new product by means of the cognitive response variable, relative preference. However, incompatibility also has a direct effect on intention to try the product. The intention to try the product in turn influences the intention to seek information about the product. Information seeking from personal sources and from impersonal sources are shown as being independent of each other. The intention to seek information then influences the intention to investigate the product. The correlations obtained in this study seem to support this type of causal linkage. Other such causal models can also be conceptualized. These causal models can be evaluated by means of path analysis (Blalock, 1974; Van De Geer, 1971) for various products and a general causal model could be developed. One interesting observation can be made from the model in Figure B with respect to causal ordering of the variables. The model indicated that

the consumer first determines her intention to buy the product and then decides to seek information or investigate the product. However, when the actual behaviors are performed the order is likely to be reversed, so that the results from investigating or information seeking will determine whether the consumer will try to product or not. The causal ordering of the behavioral intention variables, which are represented in Figure B and were included in the study, is therefore the reverse of the order in which consumers are likely to actually manifest these behaviors.

In summary, stimulus and response variables in relation to new products deserve more attention from marketers and marketing researchers. This paper has examined interrelationships between certain specific stimulus and response variables. Implications for future research have also been outlined. The major areas for further research are identification and classification of stimulus and response variables into meaningful categories, operationalization of variables, and identifying causal relationships between variables. Finally, one could extend the S-R type model and consider the moderating influences of individual characteristics and other important factors.

Appendix

New Product Descriptions

Instant Yogurt Drink: This is a new lowfat yogurt drink which tastes like a rich, creamy milkshake. It comes in four flavors: strawberry, peach, blueberry, and raspberry. The drink is packaged in 10 oz. cartons, which you only shake and pour to prepare.

Sandwich Pockets: These are flat circles of white bread which, when cut open, have pockets in the center. They can be used to make pizza, garlic bread, or a sandwich that doesn't fall apart.

Dry Soup Helper: This is a dry soup to which you add the meat. It comes in five varieties, and contains vegetables, stock, and seasonings.

Angled Toothbrush: This is a toothbrush which is constructed for most effective teeth cleaning. It has the softer bristles recommended by dentists, and a longer handle which is angled like a dental instrument, designed to reach back teeth effectively.

Multi-Surface Beautifier: This liquid restores and beautifies many surfaces, including vinyl, rubber, leather, and wood. It comes in 8 oz. plastic pump spray bottles.

Chocolate Mike Concentrate: This liquid milk additive comes in a plastic squeeze bottle and can be used to make chocolate milk. It dissolves easily in hot or cold milk.

- Wine by the Glass: This wine is packaged in 6.3 oz. bottles with a plastic drinking glass shrink-wrapped to it. It comes in four varieties--Chablis, Vin Ro se, and Burgundy.
- Boil in Bag Rice: This rice is packaged in individual plastic cooking bags. It should be boiled right in the bag and is guaranteed to cook perfectly every time. There is no need for refrigeration and it can be stored on the shelf. Cooking time is 15 minutes, but it can remain in hot water for an additional 30 minutes without becoming gummy.
- Tea for Automatic Coffee Machines: This tea is especially blended for use in automatic coffee machines. It has an extra fine texture so you get a deeper brew for making iced tea. This tea can be made in minutes in automatic machines.
- Liquid Non-Scratch Cleanser: This is a liquid household cleanser with a mild abrasive. The cleaner removes dirt like a scouring powder but it does not scratch. It can be used to clean many surfaces.
- Combination Cake & Cookie Mix: This product is a powdered mix which comes in a box and can be used to make either cake or cookies with a slight variation in the recipe.
- Detergent with Fabric Softner: This detergent has fabric softner right in it, so there is no need to add the softner during the rise cycle or in the dryer.
- Shelf Top Sealed Entrees: These entrees come packaged in a single servings in a foil pouch. To prepare, you simply boil them for five minutes, and remove from the pouch. There is no need for preservatives, and the pouches can be kept indefinitely on the shelf without refrigeration.
- Concentrated Gravy and Sauce Bars: These are refrigerated gravy and sauce bars which can be heated to form the basis for any sauce or gravy. They come in three flavors: Beef, Chicken, and White Sauce. You just add the bar to meat drippings for home-made gravy.
- Fried Chicken Maker: This is a new appliance which prepares fried chicken that tastes like chicken from a carry-out restaurant. The appliance is a pressure fryer which prepares chicken in about 15 minutes and comes in 4 or 6 quart sizes.
- Low Calorie Frozen Cakes: These frozen cakes have one-third fewer calories than others and contain no artificial sweetners. They come in three flavors: yellow with chocolate icing, chocolate with chocolate icing, banana.
- Turnable Teeth Comb: This comb has nickel-silver teeth which are designed to rotate through and de-tangle wet or dry hair. The comb also massages the scalp.
- Toaster Fries: These are frozen potato sticks similar to french fries. They come in square sheets of eight sticks which are cooked in the toaster and then broken apart to form potato sticks.
- Campagne Like Soft Drink: A soft drink with the taste and effervescence of campagne. It has no alcohol content.
- Coffee & Grain Instant Beverage: This is an instant coffee and grain beverage, made of coffee, wheat, and molasses. It is designed to look and taste like coffee.
- Blalock, H. M., Jr. Causal Inferences in Nonexperimental Research, University of North Carolina Press, 1964.
- Raju, P. S. "Theoretical Perspectives on Exploratory Behavior: A Review and Examination of Their Relevance for Consumer Research," Paper #67, Working Series in Marketing Research, The Pennsylvania State University, October, 1977.
- Robertson, T. S. Consumer Behavior, Glenview, Illinois: Scott, Foresman and Co., 1972.
- Roselius, T. "Consumer Rankings of Risk Reduction Methods," Journal of Marketing, Vol. 35 (January 1971), pp. 56-61.
- Venkatesan, M. "Cognitive Consistency and Novelty Seeking." In S. Ward and T. S. Robertson (Eds.), Consumer Behavior: Theoretical Sources, Englewood Cliffs, New Jersey: Prentice-Hall, 1973.
- Van De Geer, J. P. Introduction to Multivariate Analysis for Social Sciences, San Francisco, California: W. H. Freeman and Co., 1971.

#### References

- Bauer, R. A. "Consumer Behavior as Risk Taking." In R. S. Hancock (Ed.), Dynamic Marketing for a Changing World. Chicago, Illinois: American Marketing Association, 1960.
- Berlyne, D. E. Conflict, Arousal, and Curiosity, New York: McGraw-Hill, 1960.

# THE EFFECTS OF CHOICE COMPLEXITY AND DECISION FREEDOM ON CONSUMER CHOICE BEHAVIOR

John R. Walton, University of Minnesota  
Eric N. Berkowitz, University of Minnesota

## Abstract

The applicability of decision freedom and complexity decision concepts to a consumer soft drink choice problem was assessed. The number of choice alternatives and the quantity of soft drink were manipulated in three treatment conditions. Results indicated the amount consumed and time required to make the decision varied across these treatment conditions.

alternatives and then consume as much of that alternative as they wished. Individual subject preferences toward each alternative were not assessed beforehand. Results showed that perceived decision freedom was higher and consumption of the chosen alternative was greater in the four brand condition.

## Introduction

To an increasing degree psychological theories of decision making have been applied to consumer choice problems. Without question, these applications have expanded our knowledge about consumer choice behavior (Hansen, 1976).

At the same time, it should be recognized that consumer decision making issues (e.g., amount of consumption) differ from typical psychological research concerns. Therefore, consumer researchers should be interested in synthesizing and extending these psychological approaches to particular consumption problems. The purpose of this research is to determine the applicability of two psychological decision concepts to a particular consumer choice problem.

## Objectives and Methodology

The objective of this study is to attempt to replicate and extend these earlier findings. Specifically, it is unlikely in all choice situations that more alternatives will result in greater consumption. Furthermore, the presentation of equally valued alternatives (either two or four) and/or the quantity of product available for consumption may have some impact on choice behavior. These issues are addressed in the methodology described below.

Students in an introductory marketing class at a large midwestern university were pretested to determine their preferences toward eight different flavors of a nationally known soft drink mix. These preferences were measured on a three point scale (dislike, neutral, like).

## Background Literature

As Engel, Kollat, and Blackwell (1978) have pointed out, consumer decision making may be viewed as a process. Process elements include problem recognition, search, alternative evaluation, choice, and post-choice evaluation. This research is concerned with the concepts of choice set complexity and decision freedom as they relate to the alternative evaluation stage.

At this stage it is assumed that the individual consumer has identified a choice set consisting of  $N$  objects. Hendrick, Mills, and Kiesler (1968) used the concept of choice set complexity to explain the amount of time the individual will take to choose one of these  $N$  alternatives. Specifically, these authors hypothesized that the complexity of the choice set increases as the number of equally valued alternatives increases and/or the number of attributes used to describe those alternatives increases. The authors predicted an inverted U relationship between complexity and time to decision. The results of the study supported the prediction.

Decision freedom refers to the perception of personal volition on the part of the decision maker. Steiner (1970) asserted that decision freedom increases as the number of equally desirable alternatives in the choice set increases.

Neither the complexity nor the decision freedom explanation, however, makes specific predictions concerning the actual consumption of the chosen alternative. Reibstein, Youngblood, and Fromkin (1975) have extended the decision freedom concept to this issue.

These authors presented undergraduate students with either two brands or four brands of soft drinks. In each condition subjects were instructed to choose one of the

Approximately one week later subjects reported individually to a laboratory for the second phase of the study. Subjects were randomly assigned to one of three treatment conditions. Condition one consisted of exposure to two flavors of the soft drink with one two-quart pitcher of each flavor. Condition two consisted of two flavors of the soft drink with two two-quart pitchers of each flavor. Finally, condition three consisted of four flavors of the soft drink with one two-quart pitcher of each flavor. Pretest data allowed for the construction of equally preferred alternative sets for each individual.

The experimenter prepared the alternative set before the subject arrived. The pitchers of soft drink were placed on a table directly in front of a chair on which the subject would sit. A partition was placed around the area so that the table could not be seen by the subject until the experimenter was ready.

Upon arrival the subject's current level of thirst was measured. The subject was then informed of the nature of their choice task. They were told that as soon as they chose a flavor they should indicate that choice to the experimenter. Further, they were informed that they could drink as much of that flavor as they liked.

As soon as the subject was seated, the experimenter started a stopwatch in her pocket. When the subject announced his/her choice, the watch was stopped. After the subject had consumed as much as he/she wished, they were asked to complete a questionnaire measuring decision freedom and their perception of the purpose of the experiment. After the subject left, the experimenter measured the amount of liquid that had been poured from the pitcher. Since, in every case, all liquid poured had been drunk, this measure represented the individual's total consumption.



After reviewing the questionnaire item concerning the experimental purpose, it was clear that no subject was aware of the true nature of the experiment. After all experimental trials were concluded subjects were debriefed in class.

To summarize, dependent variables in this study include seconds to decision, perceived decision freedom, and amount of soft drink consumed. In the case of each of these dependent variables, the null hypothesis predicts no differences among the three treatment conditions. One way analysis of variance was used to test these hypotheses. The results of these analyses are described in the following section.

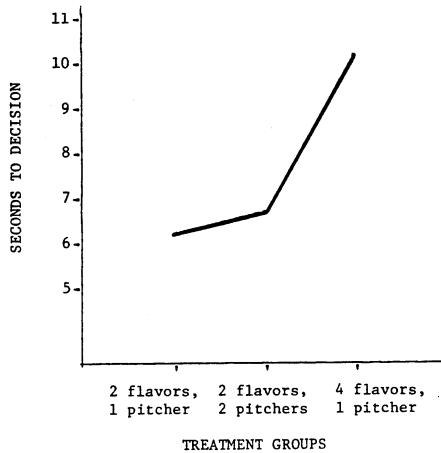
### Results

The choice complexity explanation predicts that for this relatively simple choice situation, seconds to decision should be greater in the four alternative treatment condition. Furthermore, no differences between the two alternative, one pitcher and two alternative, two pitcher conditions should be evidenced.

As shown in Figure 1, the results support these expectations [ $F(2,67) = 6.20, p < .01$ ]. Specifically, the Newman-Keuls test indicated that mean seconds to decision in the four alternative case ( $\bar{X} = 10.15$  sec.) was significantly higher than the two alternative, one pitcher ( $\bar{X} = 6.17$  sec.) and two alternative, two pitcher ( $\bar{X} = 6.78$  sec.) treatments. No difference between the two alternative conditions was observed.

FIGURE 1

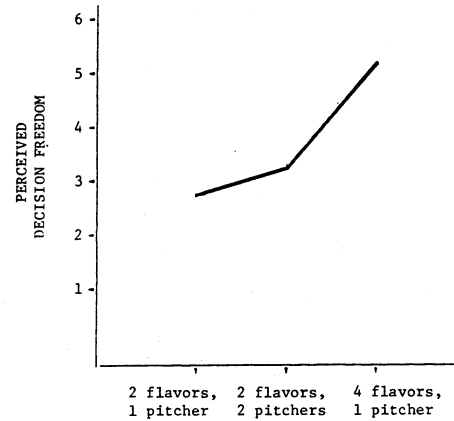
MEAN SECONDS TO DECISION FOR THE THREE TREATMENT GROUPS



The degree of perceived decision freedom also varied across treatment conditions [ $F(2,69) = 18.29, p < .0001$ ]. The highest level of perceived decision freedom was observed for the four alternative condition ( $\bar{X} = 5.17$ ). This level of decision freedom was significantly greater than found for the two alternative, one pitcher ( $\bar{X} = 2.78$ ) and two alternative, two pitcher ( $\bar{X} = 3.21$ ) treatments. Figure 2 presents a graphical portrayal of these results.

FIGURE 2

MEAN PERCEIVED DECISION FREEDOM FOR THE THREE TREATMENT GROUPS



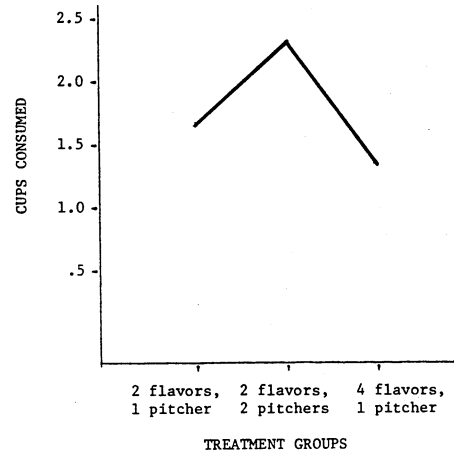
The finding that perceived decision freedom is greater for four alternatives rather than two supports the results of Reibstein et.al. Subsequently, these authors found that a larger amount of soft drink was consumed in the four alternative choice situation. An attempt to replicate this relationship between decision freedom and amount consumed was tested.

Furthermore, a quantity effect was also tested. Treatments 2 and 3 (two alternatives, two pitchers and four alternatives, one pitcher) exposed subjects to twice as much soft drink as the remaining treatment group. Significantly greater consumption in treatments 2 and 3 would support this quantity explanation.

Mean consumption levels for each treatment are depicted in Figure 3. While significant differences in amount consumed were found [ $F(2,67) = 4.07, p < .03$ ], neither the decision freedom effect nor the quantity effect was observed. Specifically, the highest consumption was found in the two alternative, two pitcher condition ( $\bar{X} = 2.3$  cups). The Newman-Keuls test indicated that this level of consumption was higher than was observed in the four alternative, one pitcher treatment ( $\bar{X} = 1.39$  cups), but not in the two alternative, one pitcher treatment ( $\bar{X} = 1.64$  cups).

FIGURE 3

MEAN CONSUMPTION OF SOFT DRINK FOR THE THREE TREATMENT GROUPS



It was expected that random assignment would preclude thirst difference between groups. However, a significant thirst difference was observed [ $F(2,69) = 4.68$ ,  $p < .02$ ]. Therefore, the analysis of variance using amount consumed as the dependent variable was rerun with level of thirst as a covariate. Significant differences among the treatments were still observed [ $F(2,66) = 3.30$ ,  $p < .05$ ]. Furthermore, the direction of those differences were as previously described.

#### Discussion

The applicability of decision freedom and complexity decision concepts to a particular consumer food choice problem was assessed. The results indicated that direct application of these concepts to consumer decision problems may not be warranted.

Specifically, the findings of Reibstein et.al., that more soft drink would be consumed when perceived decision freedom was greater, was not supported. Indeed, the opposite result was found.

Results were more promising with the complexity construct. Significantly less decision time was required in both of the two alternative conditions. Furthermore, since more soft drink was consumed in these conditions, a relationship between complexity and amount consumed is tenable.

The approach taken in this research may be appropriate for a variety of consumer food decisions. The decision process of the out-of-home food consumer is one example. The consumer makes one decision in selecting a particular restaurant or fast food outlet. A second decision is required in choosing one alternative from the range of alternatives offered on the menu.

This study suggests that the manner in which information about these alternatives is presented may impact on amount ordered (and, therefore, consumed). Presumably appropriate psychological concepts such as conflict and decision freedom will help restaurant management ascertain the most effective manner to present the alternatives to the consumer. Further research on these concepts and others will be necessary to determine the specific nature of this environment.

#### References

James F. Engel, Roger D. Blackwell, and David T. Kollat, Consumer Behavior, 3rd ed., (New York: Holt, Rinehart and Winston, 1978).

Flemming Hansen, "Psychological Theories of Consumer Choice," Journal of Consumer Research, 3(December, 1976), 117-142.

Clyde Hendrick, Judson Mills, and Charles A. Kiesler, "Decision Time as a Function of the Number and Complexity of Equally Attractive Alternatives," Journal of Personality and Social Psychology, 8(June 1968), 313-318.

David J. Reibstein, Stuart A. Youngblood, and Howard L. Fromkin, "Number of Choices and Perceived Decision Freedom as a Determinant of Satisfaction and Consumer Behavior," Journal of Applied Psychology, 60(August 1975), 434-437.

Ivan D. Steiner, "Decision Freedom," in L. Berkowitz (Ed.), Advances in Experimental Social Psychology, Volume 5, (New York: Academic Press, 1970).

James W. Harvey, University of Maryland

## Abstract

Much research is currently devoted to improved understanding of information search and processing. Much of this literature focuses on information load, information format, attention, memory, information integration schema, etc. The present paper, however, focuses on two neglected aspects of external search behavior: (a) understanding of why people search for information, and (b) the cognitive structural conditions which mediate both level and type of external search.

## Introduction

Nearly a decade of research attention to information search and processing behavior has produced a significant literature stream (e.g., Bettman, 1970; Newell and Simon, 1972; Wright, 1973; Simon, 1974; Bettman and Jacoby, 1975; Wilke, 1975; Bettman, 1977; Olson, 1977; Scammon, 1977; Olson, 1978). Two major limitations to this research exist, however. First, inadequate conceptual and empirical attention has been directed at the role external search behavior plays in the overall decision process. Most of the studies in this research tradition examine subprocesses without sufficient attention to larger processes. Second, these studies tend to emphasize the "human limits" to information acquisition (i.e., overload and format) rather than the psychological and motivational situations which may "signal" the individual to search for more or less information (e.g., Burnkrant, 1976; Harvey, 1977). Clearly, the relevance of the "quantity" and "format" types of research is understood more fully when one considers that frequently the individual may perceive little need to search externally for information. The consumer welfare and decision quality implications of this observation seem obvious (Jacoby, 1977).

Therefore, the purpose of this paper is to explore, conceptually, a hypothetical construct, Evaluative Conflict, which is postulated to mediate external information search behavior. Care is taken to integrate this search into a significant decision process conceptualization, adoption theory, as well as other related literature streams, namely cognitive structure and attitude formation. The major thrust of this work, however, is to discuss the nature of Evaluative Conflict, its underlying determinants, modes of conflict resolution, and of primary importance, how Evaluative Conflict provides a "signal" to the individual to search, externally for information.

## Adoption Theory

Adoption theory provides rich insight into the likelihood of success or failure of new products, services, and ideas (Rogers, 1962; Rogers and Shoemaker, 1971). High product failure rates persist, which contributes to renewed interest in and challenges for adoption theory (Rogers, 1976; Crawford, 1977).

<sup>1</sup> Preparation of this paper was supported, in part, by a grant from the General Research Board of the University of Maryland and has benefited from the comments of Jerry C. Olson.

Theories of adoption are generally concerned with the behavioral, social structural, information search, and cognitive processes in which the individual engages as one psychologically moves toward acceptance or rejection of an innovation. The overall process of adoption is generally conceptualized to subsume a number of other distinct and measurable subprocesses or "stages," each one of which, presumably, the decision maker "passes through." Robertson (1971) indicates that eight such subprocesses combine to form the overall adoption process: (a) problem recognition, (b) awareness, (c) comprehension, (d) attitude, (e) legitimation, (f) trial, (g) adoption, and (h) dissonance.

While much of adoption theory is devoted to conceptual and empirical examinations of the various stages which comprise the overall process, another, less well known aspect of this literature stream is one which emphasizes the dimensions which define, perceptually, any innovation. For example, Rogers (1962) contends that five perceptual dimensions exist, each of which presumably mediate the likelihood and rate of acceptance of an innovation: (a) relative advantage--the extent to which the innovation is perceived to be superior to what it supercedes or competes against; (b) compatibility--the consistency the innovation is perceived to have with established norms and values; (c) complexity--the degree of difficulty in understanding or using the innovation; (d) divisibility--the extent to which the innovation can be tried on a limited basis; and (e) communicability--the extent to which the innovation can be easily and clearly discussed with others. Each of these perceptual dimensions can be represented, however, in cognitive structure. Therefore, significant insights into acceptance of innovations is offered by such a cognitive structural orientation.

## Cognitive Structure Conceptualization of an Innovation

Two important theoretical affirmations of this paper are: first, that any innovation can be cognitively represented by beliefs (Fishbein, 1963) and that the resultant cognitive structure mediates behavioral response (McGuire, 1964; Wright, 1974). Any product, service or idea can be viewed as possessing attributes or related concepts which serve to define that stimulus for a consumer. Individual subjective feelings of association of attributes or concepts to the innovation are called beliefs. Each of the attribute beliefs within an individual's cognitive structure is either positively, negatively, or neutrally valued.

From a cognitive structure perspective, the decision process, or more precisely, the formation of a strong purchase intention (either positive or negative), may be thought of as a process of determining and comparing the presence and/or absence of both positive and negative aspects of the product. Since most innovations are perceived to possess both positively- and negatively-valued attributes, which are represented cognitively as positively- and negatively-valued beliefs, the process leading to trial is based on finding "adequate" positive reasons for purchase while discounting, depreciating, or disparaging the negative aspects of purchase (e.g., cf. Bilkey, 1951, 1953; Kotler, 1976, p. 86).

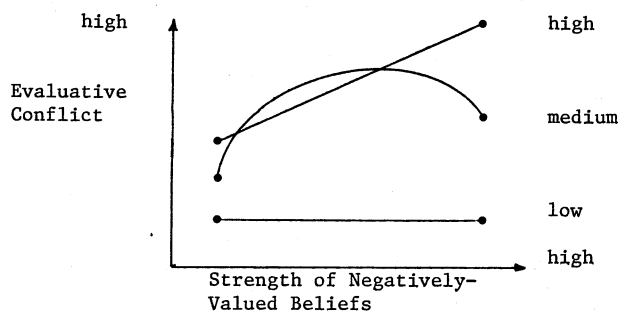
## Evaluative Conflict in Choice Processes

When the decision maker is confronted with a choice situation perceived to contain both positive and negative attributes, that person can be thought of as being in a situation of Evaluative Conflict brought about by perceived evaluative inconsistency (Fishbein and Ajzen, 1975, p. 144). That is, because both positive and negative attributes are believed linked to the same object, the individual is faced with the problem of reconciling the inconsistent evaluative (positive-negative) components of one's cognitive structure.

As summarized in Figure 1, the magnitude of Evaluative Conflict present in a given decision process is postulated to be a function of the number and strength of the positively- and negatively-valued beliefs thought to be linked to the stimulus object--i.e., a function of the level of perceived evaluative inconsistency. Maximal Evaluative Conflict is predicted when the strengths of these two categories of beliefs (a) approach equality (b) at a relatively high level. Therefore the level of Evaluative Conflict can be considered to be less than maximal whenever the strengths of either the positively- or negatively-valued beliefs differ in absolute value or whenever their strengths are low. Low overall strengths of beliefs would result in an overall feeling of indifference toward the object. Note carefully, that this discussion is based on the individual's perceived overall strength of the good and bad, not the number of beliefs.

FIGURE 1

PREDICTED RELATIONS BETWEEN THE STRENGTHS OF POSITIVELY- AND NEGATIVELY-VALUED BELIEFS AND EVALUATIVE CONFLICT



— Strengths of Positively-Valued Beliefs

Ample evidence exists that small quantities of negative information is "weighted" quite heavily in decision processes (e.g., Kanouse and Janson, 1971).

This conceptual discussion suggests that the level of evaluative conflict can be measured by the formula:

$$EC = \frac{[SG + SB]W_1}{[1 + |SG - SB|]W_2} \quad (1)$$

where EC = Evaluative Conflict  
 SG = Strength of Positively-valued Beliefs  
 SB = Strength of Negatively-valued Beliefs  
 $W_1$  and  $W_2$  = Weights to be determined by empirical inquiry

While this method of measurement may have face validity no construct validity exists for the formula. Others have suggested that the ratio of the number of positive- (or negatively-) valued beliefs to the total number

of beliefs measures the "consistency" of one's attitude (Fishbein and Ajzen, 1975, p. 100).

Given that Evaluative Conflict exists at greater and lesser levels, the key question to be addressed is its effect on human behavior, vis-a-vis the adoption process. Conceptually rich literature from the tradition of cognitive consistency theory provide several key insights into this question (Abelson, Aronson, McGuire, Newcomb, Rosenberg and Tannenbaum, 1968).

## Cognitive Conflict Theory

Two scholars provide exceptionally fruitful insights into the social psychology of Evaluative Conflict (Kelman and Baron, 1968; Kelman and Baron, 1968a). The two major contributions provided by these researchers are (a) that the presence of cognitive conflict serves as a "signal" to the individual which directs one toward specific behaviors and (b) that those behaviors may provide conflict resolution strategies for the individual. The behavior of primary importance to the present research is external information search.

## Evaluative Conflict as a Behavior Signal

The position taken by the authors is that Evaluative Conflict serves as a signal to the individual that either that person must cope with or confront the evaluative inconsistency (Kelman and Baron, 1968a). While coping with the inconsistency suggests that one fairly distinguishable set of behaviors is invoked by the individual, a different set may be employed when confronting such situations. Using the terminology of a related field of inquiry, although it appears evident that the presence and absence of Evaluative Conflict represents different task environments, the modes of resolution employed (i.e., coping or confronting) by the individual may also be thought of in a similar way (Newell and Simon, 1972). The analysis of such different task environments, therefore, represents important inquires into human problem-solving behavior. Kelman and Baron (1968a) explain that while the presence of such inconsistencies serve to motivate the individual to employ one set of resolution behaviors or another, an understanding of the various modes deserves careful attention.

## Resolution Behaviors

The researchers suggest that the resolution of situations involving Evaluative Conflict may be described in terms of the two dimensions which are postulated to "define" such inconsistencies (Kelman and Baron, 1968). First, the individual can choose either to avoid or confront the inconsistency. This is referred to as the process of conflict resolution. Second, the conflict situation can either be resolved or maintained. This is referred to as the outcome of conflict resolution. That is, in the process of either avoiding or confronting the dilemma, strategies are used by the individual which either provide a basis for "living with" (maintaining) the inconsistency or a "real" resolution (confronting). The two main elements, process (avoid or confront) and outcome (resolve or maintain), serve as descriptive dimensions for the Kelman and Baron paradigm. Since both the process and outcome elements of the paradigm have two possible subelements (avoidance or confrontation for process, and reduction or maintenance for outcome), the researchers' notion of conflict resolution can be expressed as a four-cell matrix, as shown in Table 1. The strategies listed in each cell represent potentially measurable behaviors which presumably exemplify the process and outcome combination in question. The following summary of the Kelman and Baron framework is devoted to explaining the meaning of the various strategies

(reduction, maintenance, avoidance, and confrontation combinations) which may occur in resolving the existence of Evaluative Conflict.

TABLE 1

CONFLICT RESOLUTION PARADIGM  
(ADAPTED FROM KELMAN AND BARON, 1968)

OUTCOME	PROCESS	
	Avoidance of Inconsistency	Confrontation of Inconsistency
Inconsistency Reduction	Denial	Change in attitude
	Distortion	Change in action
	Rationalization	Change in standard
Inconsistency Maintenance	<u>Derogation of source</u> Compartmentalization	<u>Influence attempt</u> Bolstering
	Institutionalized insulation	Differentiation
	Compensatory ritualism	Transcendence

Kelman and Baron explain that inconsistency reduction (the first row of Table 1) involves a refusal on the part of the individual to accept or tolerate the inconsistency with which he is presented. Inconsistency reduction processes may occur in one or two forms--first, the individual may avoid (column 1, Table 1) the inconsistency by perceiving the situation such that the inconsistency is not recognized. Alternatively, the individual may confront (column 2, Table 1) the inconsistency and then eliminate it by changing one or both of the incompatible elements.

Kelman and Baron explain that inconsistency maintenance (the second row of Table 1) is based on the proposition that the individual may be able to deal with the inconsistency by reducing the tension from the inconsistency without, however, eliminating the inconsistency itself. Inconsistency maintenance processes may take one of two forms--first, the individual may avoid (column 1, Table 1) the inconsistency by maintaining "mental distances" between the discrepant element, thus avoiding a "sharp" or direct confrontation. Second, the individual may confront (column 2, Table 1) the inconsistency and cognitively restructure the mental context within which the inconsistent elements are "placed."

In summary, Kelman and Baron suggest that four broad types of strategies exist to deal with cognitive inconsistency: (a) reduction/avoidance--where the inconsistency is "dealt" with by a perceptual recoding of the inconsistent elements; (b) reduction/confrontation--where the inconsistency is actively confronted by changing one or both of the inconsistent elements; (c) maintenance/avoidance--where the inconsistency is "dealt" with by keeping "mental distances" between the discrepant elements; and (d) maintenance/confrontation--where the discrepancy is confronted by cognitively restructuring the context of the inconsistency. The resulting four-cell matrix serves as a conceptual taxonomy of behavioral strategies for dealing with cognitive inconsistency. No additional major conceptual insights are provided by a discussion of the definitions of the various specific strategies within each cell; thus, these are omitted. The interested reader will refer to the original source.

Logically, the next major issue concerns the conditions under which specific inconsistency resolution strategies are likely to occur. For this, Kelman and Baron suggest that goals analyses provide the needed insight. That is, depending on whether the inconsistent beliefs are linked to the same or different goals, either reduction or maintenance outcomes will occur.

A single-goal situation is one in which the discrepant elements are inseparably tied to the same goal state. In this situation, since the elements cannot be "dislodged" from the same goal, Kelman and Baron suggest that reduction-type outcomes (i.e., "real" reduction in conflict) are likely to occur.

The two-goal situation is one in which the inconsistency is linked to two different goals. Since these conflict situations are ones where "dislodging" the inconsistent elements may be done more easily, the researchers suggest that maintenance-type outcomes (i.e., "living with" conflict) are likely to occur.

Concerning the process dimension of the conflict reduction paradigm, the remaining question is whether avoidance or confrontation-type strategies will occur. On this issue, Kelman and Baron offer little basis for prediction. First, the authors contend that the occurrence of either avoidance or confrontation-type processes is largely determined by the principle of "least effort." For example, some inconsistencies are relatively easy to deny or misperceive because the inconsistent element is associated with a "weak" or ambiguous stimulus. Similarly, information from a source perceived to lack knowledge, social status or attractiveness may be easily derogated. In other situations, the individual may be more sharply confronted with the inconsistency and be "forced" to grapple with it more actively and directly.

A second insight into the issue of process is offered by Kelman and Baron. To the extent the inconsistency raises questions about the achievement of short-term goals for the decision maker and the person is concerned about preserving the status quo, avoidance-type processes are likely to occur. To the extent that the inconsistency raises questions about the achievement of long-term goals and the person is concerned with preparing for future actions and interactions, confrontation-type process are likely to occur. The antecedent conditions suggested by Kelman and Baron are summarized in Table 2.

TABLE 2

ANTECEDENTS OF EVALUATIVE CONFLICT RESOLUTION  
(ADAPTED FROM KELMAN AND BARON, 1968a)

Outcome Antecedents	Process Antecedents	
	Short-run Goals	Long-run Goals
Single-Goal	Peripheral Values	Central Values
		Future Opportunities
	Reduction/Avoidance	Reduction/Confrontation
Dual-Goal	Maintenance/Avoidance	Maintenance/Confrontation

Evaluation Conflict in Adoption Theory

Because of considerable financial, performance, social and safety risk perceived to be associated with most innovations, the problem-solving task for these products, services or ideas can likely be categorized as single-goal decisions which raise questions about one's central values and future opportunities. Seemingly, one could

label such task environments as being characterized by high ego-involvement. Therefore, Reduction/Confrontation resolution behaviors can be expected to dominate decision processes for innovations. The issue of significance to management and public policy decision makers is a clear understanding of the nature of this resolution strategy. To achieve this insight, attention is directed to the cognitive structural conceptualization of an innovation proposed earlier in this paper.

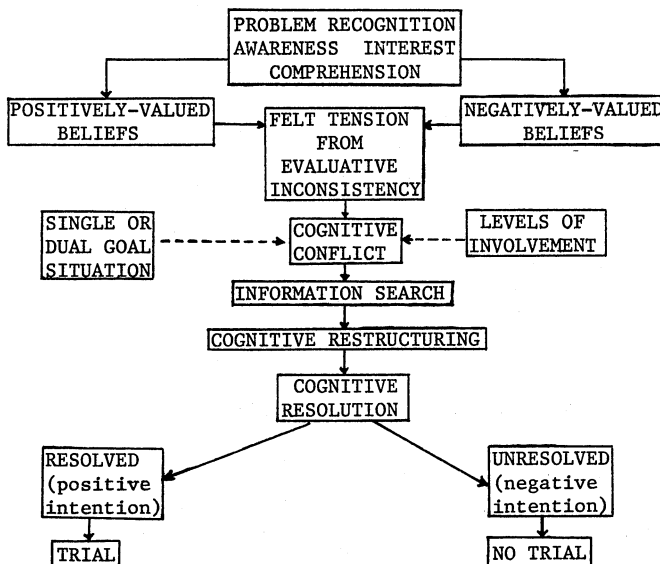
The decision-making process for an innovation which is perceived to possess both positively- and negatively-valued beliefs will result in a psychological situation of high Evaluative Conflict. This state will "signal" the individual, therefore to resolve the conflict. This signal will serve to mediate external information search since memory sources would already be "accounted for" in the perception of levels of evaluative inconsistency. Consequently, high levels of Evaluative Conflict will lead to high levels of information search. Furthermore, this information search will include larger information search from trusted sources (i.e., non-propagandistic) since the perceived negativity in the decision process will also result in high perceived risk. This process is summarized in Figure 2.

The significance of this conceptual development is clear. Improved understanding of why decision makers search for information and the cognitive structural conditions which mediate the level and type of search provide important insight into marketing management and public policy decision making. With this understanding, for example, prospective adoptors of innovations can be induced to search for more or less information. Furthermore, the need for more or less trusted sources of information can be predicted.

Once a cognitive structural measure of Evaluative Conflict is accomplished for a specific innovation, analysis of that structure will offer insight into level and type of information search. Through this analysis, traditional communication programs can be created to alter this cognitive structure which will lead to a modification of Evaluative Conflict and therefore, a directly proportional need to search for information from external sources.

FIGURE 2

A SCHEMATIC REPRESENTATION OF EVALUATIVE CONFLICT AROUSAL AND RESOLUTION EFFECTS ON THE FORMATION OF INTENTIONS TO TRY AN INNOVATION



## References

- R. P. Abelson, et al., Theories of Cognitive Consistency: A Sourcebook, Chicago: Rand McNally, 1968.
- J. R. Bettman, "Information Processing Models of Consumer Behavior," Journal of Marketing Research 7 (1970), 370-6.
- \_\_\_\_\_, "Consumer Information Acquisition and Search Strategies," (unpublished working paper), University of California, Los Angeles, 1977.
- J. R. Bettman, and J. Jacoby, "Patterns of Processing in Consumer Information Acquisition," Center for Marketing Series Working Paper Series, No. 31 University of California at Los Angeles, Graduate School of Management, 1975.
- W. J. Bilkey, "The Vector Hypothesis of Consumer Behavior Analysis," Journal of Marketing 3 (1953): 18-25.
- Robert E. Burnkrant, "A Motivational Model of Information Processing Intensity," Journal of Consumer Research, Vol. 3 (June 1976), 21-30.
- C. M. Crawford, "Marketing Research and the New Product Failure Rate," Journal of Marketing, 41 (April 1977), 51-61.
- M. Fishbein, "An Investigation of the Relationships Between Beliefs About an Object and the Attitude Toward the Object." Human Relations 16 (1963: 233-40.
- M. Fishbein, and I. Ajzen, Belief, Attitude, Intention and Behavior. Reading, Massachusetts: Addison-Wesley, 1975.
- James W. Harvey, "The Legitimation Stage in the Process of Adoption of Innovations," (unpublished doctoral dissertation), Department of Marketing, Pennsylvania State University, August, 1977.
- J. Jacoby, "Information Load and Decision Quality: Some Contested Issues," Journal of Marketing Research, Vol. XIV (November 1977), 569-73.
- D. E. Kanouse, and L. R. Janson, Jr., "Negativity in Evaluations." In Attribution: Perceiving the Causes of Behavior, pp. 47-62. Edited by E. E. Jones et al. Morristown, New Jersey, 1971.
- H. C. Kelman, and R. M. Baron, "Determinants of Modes of Resolving Inconsistency Dilemmas: A Functional Analysis." In Theories of Cognitive Consistency: A Sourcebook, pp. 670-83. Edited by R. P. Abelson, et al. Chicago: Rand-McNally, 1968.
- H. C. Kelman, and R. M. Baron, "Inconsistency as a Psychological Signal." In Theories of Cognitive Consistency: A Sourcebook, pp. 331-6. Edited by R. P. Abelson, et al. Chicago: Rand-McNally, 1968a.
- P. Kotler, Marketing Management: Analysis, Planning and Control, Third Edition. Englewood Cliffs, New Jersey: Prentice-Hall, 1976.
- W. J. McGuire, "Inducing Resistance and Persuasion: Some Contemporary Approaches," In Advances in Experimental Social Psychology, Edited by L. Berkowitz. Vol. 1. New York: Academic Press, 1964, 121-229.
- A. Newell and H. A. Simon, Human Problem Solving. Englewood Cliffs, New Jersey: Prentice-Hall, 1972.

J. C. Olson, "Theories of Information Encoding and Storage: Implications for Consumer Research." Paper presented at the joint Carnegie-Mellon University--American Marketing Association Conference "The Effect of Information on Consumer and Market Behavior," May, 1977.

J. C. Olson, "Inferential Belief Formation in the Cue Utilization Process," Advances in Consumer Research, Vol. 5, 1978, 706-13.

T. S. Robertson, Innovative Behavior and Communications. New York: Holt, Rinehart and Winston, 1971.

E. M. Rogers, Diffusion of Innovations. New York: The Free Press of Glencoe, 1962.

E. M. Rogers, "New Product Adoption and Diffusion," Journal of Consumer Research, Vol 2 (March 1976), 290-301.

E. M. Rogers and F. F. Shoemaker, Communication of Innovations. New York: The Free Press, 1971. Consumer Research, Vol. 4 (December 1977), 148-55.

H. A. Simon, "How Big Is A Chunk?" Science, 183 (February 1974), 482-8.

William L. Wilke, "How Consumers Use Product Information," National Science Foundation (RANN) Report, 1975.

P. Wright, "The Choice of a Choice Strategy: Simplifying Versus Optimizing," Faculty Working Paper, No. 163. University of Illinois: College of Commerce and Business Administration, 1973.

P. L. Wright, "The Cognitive Processes Mediating Acceptance of Advertising," Journal of Marketing Research Vol. X, (February, 1974), 53-62.

James R. Bettman, University of California, Los Angeles

## Abstract

This paper discusses three studies on consumer decision making. Issues of construct definition, construct measurement, and theory development are addressed, and several priority areas for research on consumer choice are considered.

## Introduction

The three papers presented in this session deal with the broad area of consumer decision making. Each examines some aspect of the choice process and attempts to add to our knowledge of that process. In the following, each paper is briefly discussed. Then implications for future research on consumer choice are considered.

## Evaluative Conflict and Information Search in the Adoption Process

Harvey's paper attempts to provide a framework for examining one determinant of external information search, namely the level of evaluative conflict felt by the consumer. This area is an important aspect of the consumer choice process, and search for information and conflict or related variables (e.g., arousal, incongruity) are core constructs in several consumer choice theories (Howard and Sheth, 1969; Hansen, 1972, Howard, 1977; Bettman, 1978). In addition, the framework proposed by Kelman and Baron (1968) and presented by Harvey appears to offer some interesting potential insights into how consumers respond to conflict. In particular, variables which may affect individuals' conflict reaction strategies are presented: the type and number of goals active in a choice situation, and the degree of effort required to execute a strategy, for example.

The contributions of the paper remain potential, however, due to some unresolved problems. First, the construct of evaluative conflict (EC) is not yet well delineated or defined, a problem common to many studies of consumer choice using fairly complex and abstract constructs. Prior work on specifying conflict is not addressed (e.g., Berlyne, 1957). This lack of rigorous specification makes measurement of such constructs a very difficult task. In the present paper, an intuitive measure for EC is proposed, but its goodness cannot be evaluated without further specification of the construct being measured. I would simply comment that the measure presented in Formula 1 appears to require ratio scales for SG and SB. This imposes a strong requirement on the researcher, to say the least. Also, it is not clear why the case of high positive and low negative beliefs should entail moderate conflict, as Formula 1 and Figure 1 seem to imply, rather than low conflict.

A second general concern is that there are no specific theoretical propositions advanced which could form the basis for future research. A model which depicted the relationship of several constructs to degree and direction of external search could be highly useful. Harvey's paper hints at such a model, but does not specify one in any detail. In building such a model, I would suggest that such notions as the perceived costs and benefits of search; choice environment factors (such as availability of information or time pressure); and individual differences (abilities, concern with choice optimality) be included. The focus on conflict alone seems too narrow (Bettman, 1978, Chapter 5).

Finally I would quarrel with Harvey on a more specific point, whether decisions about most innovations are single-goal decisions. Harvey claims they are, and I find this hard to accept. Harvey himself notes that there are financial, performance, and safety factors associated with innovations, and these are likely to have corresponding goals. Thus I believe that such choices are more likely to be multiple-goal decisions.

In summary, the Harvey paper alludes to some interesting vistas for research, and I would encourage him and others to flesh out some of these areas more fully.

## Stimulus-Response Variables in New Product Research

The Raju paper explores how consumers' perceptions of innovations are related to their reactions to those innovations. Again, this is an issue which is important for understanding consumer choice, as the perceptual encoding of the choice alternatives can strongly influence how those alternatives are evaluated and compared. The types of stimulus variables proposed by Raju may provide insight into variety seeking behavior, one of the toughest and most underresearched areas in consumer choice.

In attempting to implement the study, however, some serious problems arose. First, there seem to be substantial measurement problems. Single item scales are used to measure relatively complex constructs. There is no evidence that these scales correspond to the constructs they are meant to represent; in part, this difficulty is compounded because the constructs themselves are not well-specified. The so-called behavioral response measures exemplify this problem. These measures are hypothetical cognitive responses, not behavioral measures. I also have some difficulty in determining how to differentiate investigatory and seeking responses, for example. If one picks up a package to examine it out of curiosity and looks at price and ingredients while so doing, it is not clear whether that is seeking or investigating. In addition, the "cognitive" response measure, preference, is actually an affective measure. Thus, more careful specification of constructs is needed.

Secondly, I believe the study could have been made stronger by specifying a theory a priori and manipulating the independent variables contained in that theory. Previous research on diffusion of innovations (Rogers and Shoemaker, 1971) and on the impact of consumer perceptions on innovative behavior (Ostlund, 1974) showed the impact of compatibility, for example. There would appear to be enough information from this prior research to start with a theoretical causal model and test that model. I have other problems with the study (e.g., I see the paper as using an S-O-R rather than S-R model, since perceptions of stimuli are the focus), but the two above are my major concerns. Despite these issues, I believe variety seeking is a neglected area of consumer research, and hope that Raju's exploratory efforts will stimulate other research on this topic.

## The Effects of Choice Complexity and Decision Freedom on Consumer Choice Behavior

The paper by Walton and Berkowitz examines some aspects



of consumer choice which I feel will be of growing importance in future research: decision time and consumers' perceived freedom of choice. Examination of factors influencing choice times can lead to insights into the heuristics used in making choices and to increased knowledge about underlying choice processes (Wright, 1977; Sternberg, 1977). Consumers' perceptions of choice freedom could have important implications for consumer satisfaction research and for policy. For example, the Walton and Berkowitz study finds that for alternatives low in complexity, presenting subjects with more choice alternatives (flavors) leads to longer choice time and more perceived freedom. This might pose an interesting conflict, if shorter choice times and higher perceptions of freedom of choice are both valued by consumers. However, the very simple stimuli used make generalizations to increased numbers of brands problematic.

I have two comments on problems with the study. First, the authors, like many previous researchers, do not present the Hendrick, Mills, and Kiesler (1968) study accurately. Hendrick et.al. did not define an overall concept of choice set complexity. Rather, they noted that number of alternatives and complexity of each alternative (number of dimensions emphasized) were related to choice time. The present study thus uses number of alternatives, and not complexity as defined by Hendrick et.al. In fact, complexity in the Hendrick et.al. sense seems quite low for the Walton and Berkowitz study. In addition, Hendrick, Mills, and Kiesler's data do not show an inverted U relationship between complexity and decision time (and in fact neither Hendrick et.al. nor Walton and Berkowitz investigate such a relationship). Hendrick et.al. found that when only one dimension of the alternatives was considered (low complexity), time to choose among four attractive alternatives was significantly longer than choice time for a set of two attractive and two unattractive alternatives. When many dimensions were stressed (high complexity), these choice times were not significantly different.

A second potential problem concerns the finding that subjects drank more in the two flavor, two pitcher condition, a finding the authors attribute to complexity. However, as noted above, the authors did not vary complexity, but number of alternatives. In addition, number of different alternatives does not order the 2 flavor, 2 pitcher and 2 flavor, 1 pitcher findings (which differ in direction although the difference apparently only approaches significance). These findings seem more readily explained by noting that subjects may not feel right about drinking a lot if only one pitcher is available, but may feel more comfortable in doing so if they see two pitchers are there. This form of subject evaluation apprehension may compromise the internal validity of the experiment.

Decision time and choice freedom are worthy of further research. I would encourage examining these variables in more complex choice tasks.

#### Implications for Future Research on Consumer Decision Making

##### Construct Definition, Construct Measurement, and Theory Development

Two major issues arose in the above discussion: problems related to construct definition and measurement and theory development, and some areas of interest for future research. Each of these topics is considered below.

Construct validity appears to be a major problem in consumer choice research. In particular, many, if not most, constructs used in consumer choice research are not rigorously defined. For example, constructs such as conflict, information search, information, perceived risk, novelty, and so on are typically not carefully delimited

and differentiated from other related constructs. This lack of specificity of course makes it very difficult to determine whether a particular operationalization does or does not correspond to the construct. Such correspondence must be assessed to evaluate the strength of any test of a theory involving that construct, however (Calder, Phillips, and Tybout, 1978). Hence, one major priority for research on choice is to carefully define any constructs used in building theories. Given such construct definition, multiple measurements of each construct are desirable. Recent approaches to assessing reliability and construct validity for such multiple measurements within a structural equations framework seem highly appropriate (Bagozzi, 1978). In addition to definition and measurement of constructs, development of theory is important to progress in research on choice. Specification of particular propositions relating constructs of interest and examination of these propositions within the framework of an overall network of relationships is crucial (Bagozzi, 1977, 1978).

##### Priority Areas for Research on Consumer Choice

Some priority areas for research were noted above; conflict and conflict resolution strategies; information search; studies of factors which influence decision time; and perceptions of product stimuli. There are several other currently underresearched topics which seem extremely promising for extending our knowledge of consumer decision making processes in different directions than those just mentioned: memory; detailed analyses of consumer responses to information; uses of verbal report data on choice processes; and investigation of process measurement methods.

Memory. Research on consumer memory has been a neglected area of study. However, it has become increasingly clear that what is in consumer memory must be examined to study such phenomena as information search, encoding of stimuli, the inferences made by consumers, and so on. (Bettman, 1978; Olson, 1978b). One area of particular interest is how consumers organize information in memory, what chunks, schemas (Markus, 1977), or scripts (Abelson, 1976) are available to them. The specific inferences by consumers from product stimuli, advertising, word of mouth, and other product-related information may be heavily dependent on what data are in memory and how those data are organized (e.g., Olson, 1978a). In addition, knowledge of how memory is organized can lead to more effective presentation of information to consumers, if the organization of that information is made congruent with the organization of memory. A major research problem in this area is how to measure such constructs as schemata and scripts. In part, this problem stems from incomplete specification of these notions. For some promising attempts, see Markus (1977) and Clary, Tesser, and Downing (1978).

##### Detailed Analyses of Consumer Responses to Information.

There has been increased interest of late in detailed examination of consumer responses to communications. In particular, the cognitive response approaches introduced by Wright (1973) are being used by more researchers (e.g., Sternthal, Dholakia, and Leavitt, 1978; Edell and Mitchell, 1978; Olson, Toy, and Dover, 1978) to test theoretical notions about how persuasive communications impact consumers. One priority for research in this area, therefore, is to develop such detailed propositions for empirical test. A second promising approach might be to use cognitive response methods to study consumer response to other stimuli. For example, cognitive responses could be elicited to probe consumers' reactions to actual experiences with a product. That is, a consumer might taste or use a product and immediately thereafter be asked to list the thoughts and reactions they had during this usage experience. Such approaches might lend insights into such important and unresearched issues as

how consumers learn by forming inferences from consumption outcomes; what types of plans or instructions to themselves consumers generate based on usage experience; and how consumers integrate the information they gain from product communications and actual usage experiences.

Uses of Verbal Report Data on Choice Processes. Recently there has been some concern over what can be learned from verbal report data on choice processes. In particular, Nisbett and Wilson (1977) argue that subjects are largely unaware of their choice processes and cannot report accurately about them. This conclusion, if correct, would have serious implications for methodology in choice research (e.g., for the cognitive response research noted above). However, the Nisbett and Wilson conclusions may be overstated. Ericsson and Simon (1978a, b, c), in a masterful series of papers, outline the types of verbal data on cognitive processes that are likely to be most useful. Using an information processing framework, Ericsson and Simon argue that some internal states and intermediate steps in processing are available in subjects' memories. They argue further that subjects can accurately report these internal states and processing steps from short-term memory if concurrent protocols are taken (i.e., while the subject is performing the task of interest). Such protocols should be more accurate under general instructions to verbalize rather than under instructions to produce specific information (Ericsson and Simon, 1978a,b).

Retrospective questioning about choice processes should be more accurate when general instructions to verbalize are used (as in most cognitive response studies); when the time between process and verbalization is short; and when subjects are simply asked to report the information attended to and intermediate results they remember, rather than to report information requiring inference on the part of the subject (i.e., How did you make a choice?) "We think it very important to distinguish direct verbalization of internal states from verbalizations that are partly or primarily the results of intermediate processes like abstracting, inference, search for prototypic examples, and so on. If the aim of probing is to collect verbal data on the internal states of the cognitive processes under study, it should be the object of a verbalization procedure to keep the intermediate processes at a minimum." (Ericsson and Simon, 1978c, p. 13).

Thus, verbal report data on choice processes should be collected with general instructions to verbalize, preferably concurrently with choice, but as temporally close to choice as possible in any case. Such data might then be coded with respect to the types of information attended to, intermediate results, or other internal states that are revealed.

Process Measurement Methods. A final important area for research is study of the properties of various methods for investigating choice processes. Process methods may have non-obvious biases (e.g., see Arch, Bettman, and Kakkar (1978) for biases in information display board methods), and most process methods now available are very time-consuming and yield data which are difficult to analyze. New methods which are more efficient would be desirable. Research should also be carried out on methods such as eye movement analysis and response time analysis (Russo, 1978) to ascertain their usefulness for consumer choice research.

There is thus an array of exciting research questions on consumer choice to be investigated. By paying careful attention to construct and theory development and to appropriate use of process measurement methods, significant advances in our understanding of consumer choice processes should result.

## References

- Robert P. Abelson, "Script Processing in Attitude Formation and Decision Making," in John S. Carroll and John W. Payne, eds. Cognition and Social Behavior (Hillsdale, N. J.: Lawrence Erlbaum, 1976, 33-45).
- David C. Arch, James R. Bettman, and Pradeep Kakkar, "Subjects' Information Processing in Information Display Board Studies," in H. Keith Hunt, ed., Advances in Consumer Research, Volume 5 (Chicago: Association for Consumer Research, 1978, 555-60).
- Richard P. Bagozzi, "Structural Equation Models in Experimental Research," Journal of Marketing Research, 14 (May 1977), 209-26.
- Richard P. Bagozzi, "Reliability, Validity, and Theory Testing in Marketing Research: A Structural Equation Approach," Working paper, School of Business Administration, University of California, Berkeley, 1978.
- D.E. Berlyne, "Uncertainty and Conflict: A Point of Contact Between Information-Theory and Behavior-Theory Concepts," Psychological Review, 64 (November 1957), 329-39.
- James R. Bettman, An Information Processing Theory of Consumer Choice (Reading, Mass.: Addison-Wesley, 1978).
- Bobby J. Calder, Lynn W. Phillips, and Alice M. Tybout, "Universalistic and Particularistic Consumer Research," Working paper, Graduate School of Business, Northwestern University, 1978.
- E. Gil Clary, Abraham Tesser, and Leslie L. Downing, "Influence of a Salient Schema on Thought-induced Cognitive Change," Personality and Social Psychology Bulletin, 4 (Winter 1978), 39-43.
- Julie A. Edell and Andrew A. Mitchell, "An Information Processing Approach to Cognitive Response," in Subhash C. Jain, ed., Research Frontiers in Marketing: Dialogues and Directions (Chicago: American Marketing Association, 1978, 178-83).
- K. Anders Ericsson and Herbert A. Simon, "Protocols as Data I: Effects of Verbalization," Working paper, Department of Psychology, Carnegie-Mellon University, 1978a.
- K. Anders Ericsson and Herbert A. Simon, "Protocols as Data II: Veridicality of Verbalized Information," Working paper, Department of Psychology, Carnegie-Mellon University, 1978b.
- K. Anders Ericsson and Herbert A. Simon, "Retrospective Verbal Reports as Data," Working paper, Department of Psychology, Carnegie-Mellon University, 1978c.
- Flemming Hansen, Consumer Choice Behavior: A Cognitive Theory (New York: Free Press, 1972).
- Clyde Hendrick, Judson Mills, and Charles A. Kiesler, "Decision Time as a Function of the Number and Complexity of Equally Attractive Alternatives," Journal of Personality and Social Psychology, 8 (1968), 313-18.
- John A. Howard, Consumer Behavior: Application of Theory (New York: McGraw-Hill, 1977).
- John A. Howard and Jagdish N. Sheth, The Theory of Buyer Behavior (New York, Wiley, 1969).
- Herbert C. Kelman and Reuben M. Baron, "Determinants of Modes of Resolving Inconsistency Dilemmas: A Functional Analysis," in Robert P. Abelson et al., eds., Theories

of Cognitive Consistency: A Sourcebook (Chicago: Rand McNally, 1968, 670-83).

Hazel Markus, "Self-Schemata and Processing Information about the Self," Journal of Personality and Social Psychology, 35 (February 1977), 63-78.

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.

Jerry C. Olson, "Inferential Belief Formation in the Cue Utilization Process," in H. Keith Hunt, ed., Advances in Consumer Research, Volume 5 (Chicago: Association for Consumer Research, 1978a, 706-13).

Jerry C. Olson, "Theories of Information Encoding and Storage: Implications for Consumer Research," in Andrew A. Mitchell, ed., Effects of Information on Consumer and Market Behavior (Chicago: American Marketing Association, 1978b, in press).

Jerry C. Olson, Daniel R. Toy, and Philip A. Dover, "Mediating Effects of Cognitive Responses to Advertising on Cognitive Structure," in H. Keith Hunt, ed., Advances in Consumer Research, Volume 5 (Chicago: Association for Consumer Research, 1978, 72-78).

Lyman E. Ostlund, "Perceived Innovation Attributes as Predictors of Innovativeness," Journal of Consumer Research, 1 (September 1974), 23-29.

Everett M. Rogers and F. Floyd Shoemaker, Communication of Innovations, (New York: Free Press, 1971).

J. Edward Russo, "Eye Fixations Can Save the World: A Critical Evaluation and a Comparison between Eye Fixations and Other Information Processing Methodologies," in H. Keith Hunt, ed., Advances in Consumer Research, Volume 5 (Chicago: Association for Consumer Research, 1978, 561-70).

Robert J. Sternberg, "Component Processes in Analogical Reasoning," Psychological Review, 84 (July 1977), 353-78.

Brian Sternthal, Ruby Dholakia, and Clark Leavitt, "The Persuasive Effect of Source Credibility: Tests of Cognitive Response," Journal of Consumer Research, 4 (March 1978), 252-60.

Peter L. Wright, "The Cognitive Processes Mediating Acceptance of Advertising," Journal of Marketing Research, 10 (February 1973), 53-62.

Peter L. Wright, "Decision Times and Processes on Complex Decision Problems," Working paper, Graduate School of Business, Stanford University, 1977.

## Abstract

This paper discusses the theorizing, research designs and data presentations in three studies of decision processes. The focus is on the development of theoretical propositions consistent with prior research and on research strategies for testing such propositions.

## Introduction

Each of the three papers to be reviewed (Walton and Berkowitz, 1978; Raju, 1978) presents propositions about aspects of the decision-making process, and two papers present data. The theoretical propositions are discussed in terms of their clarity and consistency with earlier research. In discussing the empirical tests, the chosen research strategy is discussed in terms of alternate strategies one might use for researching the same question. The intent of this discussion is to position these three studies relative to the larger stream of research of which they are a part.

Walton and Berkowitz: The Effects of Choice Complexity and Decision Freedom on Consumer Choice Behavior

The Walton and Berkowitz experiment studied the relationship between two antecedent factors--- the number of available alternatives ( $N_{alt}$ ) and the amount of each alternative available for consumption--- and three behaviors---total time to decide one's preference, post-decisional feelings about one's earlier decision freedom, and actual consumption of the chosen alternative. The experiment's motivation is commendable. Variables which theorizing suggests may play roles in conditioning decision processes were selected, and a study that represents both a replication and an extension of prior experiments was conducted. Similarly, the design of the experiment seems free of obvious artifacts, as best we can tell from the report, and includes one procedural feature---the equating of the attractiveness of the stimulus alternatives presented to each subject---which helps clarify the  $N_{alt}$  manipulation.

Relationship to Prior Research

The presentation of the data might be more enlightening if it were accompanied by more detailed theorizing about hypothesized relationships and more direct interpretations of the disparity between the Walton and Berkowitz results and the earlier Reibstein et al (1975) results. Let me briefly discuss the disparity issue first.

Reibstein et al found that subjects who chose between four soft drinks reported having felt more decision freedom and drank more of the chosen alternative than subjects who chose between two drinks. Walton and Berkowitz note that the relative attractiveness of the four drinks in the Reibstein et al study was not scaled a priori, nor equated at the individual

subject's level. They correct this in their study, and hint that it is a crucial design feature, yet do not explain why explicitly. It seems desirable to make explicit why this is important in interpreting the observed disparity in findings; in so doing, I am presumably only stating what Walton and Berkowitz had in mind but did not state.

There is a greater probability that a set of four randomly chosen drinks will include a drink the subject finds especially attractive than if only two of the four are offered. Hence, with no prior scaling of the drinks' attractiveness to each subject, we would expect more of Reibstein et al's "four alternative" subjects to find a highly attractive drink in their set than is true of the "two alternative" subjects. If we then merely assume that the amount someone drinks depends on how much he likes what he is drinking, we would predict greater consumption of the chosen drink in the case where  $N_{alt}=4$ , just as Reibstein et al observed. This consumption behavior would not depend in any way on perceived choice freedom, and the interpretation in that vein offered by Reibstein et al would be questionable. Walton and Berkowitz eliminated this design artifact, and found no relationship between  $N_{alt}$  and volume consumed, even though perceived decision freedom did increase as  $N_{alt}$  increased.

Walton and Berkowitz were also interested in how  $N_{alt}$  and amount available per alternative affect total decision time. The reason behind the interest in amount per alternative is not made clear.  $N_{alt}$  has often been cited as a determinant of decision time, but it is prudent to examine the record.

Factors Affecting Decision Time

The authors refer to Hendrick, Mills, and Kiesler (1968) as support for the general proposition that decision times first increase, then decrease, as the complexity of the choice increases, with  $N_{alt}$  seen as one factor contributing to complexity. Unfortunately, the authors misconstrue the oft-cited Hendrick et al data, just as Hendrick et al and legions of others have done, a point documented at some length elsewhere (Wright, 1978). Briefly, Hendrick et al contrasted subjects facing (a) four attractive neckties or (b) two attractive and two unattractive neckties when (i) one cue was made salient or (ii) many cues were supposedly salient. Subjects chose twice, in different conditions of the design. On the first choice problem, extra attractive alternatives caused higher total decision times when one cue was salient. When many cues were salient, there was no reliable effect due to extra alternatives. In the second problem, extra attractive alternatives led to higher decision times regardless of how many cues were salient. Hence, in no case did increasing the proportion of attractive alternatives in the set cause increases in decision times. Hendrick et al stated otherwise in their conclusion: "Under the many dimension condition, decision time was shorter for four equally attractive alternatives than for two equally attractive and two unattractive alternatives (1968, p. 317)". The study is so often cited and the design or data so often mischaracterized that it deserved comment. (For more discussion of the data from Hendrick et al.

see Wright, 1978).

It also seems timely to take stock of the accumulated evidence regarding  $N_{alt}$  effects on times to decide between complex alternatives. Many researchers have proposed an inverted-U shaped relationship between decision times and a choice problem's complexity, with  $N_{alt}$  or the number of salient cues to be used for evaluations ( $N_{cue}$ ) seen as contributors to complexity. Empirical tests include Kiesler, 1966, Pollay, 1970; Jacoby, Speller, and Kohn, 1974; Jacoby, Szybillo, and Busato-Schach, 1977, and Van Raiij, 1976, in addition to Walton and Berkowitz. Often, manipulations of  $N_{alt}$  have been confounded with manipulations of the disparity between alternatives in the set, a problem Walton and Berkowitz apparently avoided. And manipulations of  $N_{cue}$  have likewise been questionable (Wright, 1978). To date, no study has produced clear evidence that increasing either  $N_{alt}$  or  $N_{cue}$  shortens someone's total decision time, except Van Raiij who used a within-subject design (which may have heightened subjects awareness of the  $N_{cue}$  manipulation). Further, even if such an effect had been empirically documented, it would not necessarily be due to increased complexity; it is quite possible to simplify someone's choice problem by adding an extra cue, providing it is a discriminator cue.

### Research Strategy

Two final comments are offered dealing with general research strategy where decision times are of interest. First, Walton and Berkowitz predict that adding more flavors to the choice set in their experiment will cause longer decision time, since this whole problem seems intuitively to be relatively simple. With a complex variable like "choice problem complexity" and a hypothetical humpbacked U-shaped relationship in mind, banking on intuition about just where one's manipulation will tap into the "complexity" continuum seems very risky. Depending on the other problem parameters which may affect complexity, an incremental increase in  $N_{alt}$  might push times up, bracket the hypothetical inflection point, or pull times down, and a researcher's intuition about where the inflection point occurs in a given setting may not be good. In any case, given the wonderful flexibility of U-shaped hypotheses in handling data, it seems desirable to offer strong reasoning or supporting empirical results to show that a predicted relationship of form X was justifiable a priori.

Second, it seems desirable to recognize that total decision time is spent performing conceptually different activities, that these activities are what is really of concern, and that a given total time prediction should be reasoned out in terms of the relevant activities. For example, in predicting that there will be no difference in total times between "two drinks, one pitcher" and "two drinks, two pitchers each", the authors apparently assumed that physical scanning time, which should be somewhat longer when there are twice as many pitchers to inspect, would be a trivial component of total decision time. This should at least be explicitly discussed. Further, is the increase in time when  $N_{alt}$  doubled in this experiment due to increased  $N_{alt}$  scanning time or increased preference analysis time or both? Spelling out the thinking about the relevant cognitive activities that may be affected by a problem parameter is desirable, as well as seeking data analysis methods that enable one to identify which activity was affected.

## Raju: Stimulus-Response Variables in New Product Research

Raju's interest is more in the line of taxonomy development. He examines the intercorrelations between a sampling of information-seeking acts, between a set of general dimensions of new products, and between these two sets of data, relying on reported perceptions about product attributes and about action intentions.

### Research Strategy

One aspect of Raju's research strategy that deserves comment is the treatment of the "action" variables. Raju was interested in the types of information seeking acts that people see as similar in judging their information search intentions. To explore this, the intercorrelations between reported intentions to perform four information-seeking acts, as evoked by numerous stimulus products, were examined. As it turned out, three of the acts (Numbers 2,3 and 5 in Table 2, Raju, 1978) were closely correlated. These three have in common that a person "happens to notice" the actual product, and this stimulates immediate information search. One might hypothesize that these acts represented minor variants of a general behavioral factor we might call "physical-product stimulated search". Note that Behavior #4, which did not correlate closely with the other three, is the only one where search is triggered by a stimulus (an ad) other than the physical product. However, this interpretation may be questioned because Behavior #4 also differed by depicting an information seeking act with a deferred payoff; the data only arrive much later. Hence, it is unclear how to interpret the pattern of intercorrelations.

How might this question of the types of information seeking acts that people see as similar or different in judging their action intent be studied? Two strategies might be used. One might begin by assembling as exhaustive a list as is possible of the specific information seeking acts or action contexts a person might envision. A large random sample of these acts, or even the entire set, might then be used as separate intent measures in a study like Raju's, where products are the stimulus.

An example of this research strategy is the numerous studies by Triandis and his colleagues (Triandis, 1972) regarding the dimensions of "interpersonal action". First, a broad list of concrete interpersonal acts was compiled, using verbs as the basis. Then subjects reacted to stimulus persons described in various ways by stating their intentions to perform the various interpersonal acts on the lists. A factor analysis identified a small set of interpersonal-action contexts to which people seem to react: Self-superior contexts, self-subordinate contexts, intimate act contexts, etc. Replications using the same "start from scratch" methodology in different cultures have yielded evidence of much cross-cultural similarity in the dimensions, although different tendencies are noted in the way specific "context-other person" combinations are evaluated.

A second approach would be to hypothesize different clusterings of acts, based on prior evidence or some rudimentary theory about what makes one act of information seeking different from another. Then one could pick a few specific acts intended to represent each of the hypothesized clusters, to be used as intent measures.

In picking concrete acts or action contexts, one would consciously strive to clearly differentiate the acts along one's hypothesized dimensions. For example, if one thought that "physical product stimulated search" differs in people's minds from "symbolic product stimulated search", and that "quickly completed acts of search" differ from "wait for the data to arrive" acts, one would create acts which keep these dimensions distinct, for inclusion in the set that elicits intent statements: for example, "notice it in a store, pick it up and look at it" (physical product stimulated/immediate data), "notice it in a vacationing friend's house/leave a note asking for advice about it when the friend returns" (physical product stimulated/delayed data), "notice an ad for it, read the ad" (symbolic product stimulated/immediate data), "notice an announcement for a brochure, send for the brochure" (symbolic product/delayed data). The author and Alain Cousineau tried this approach several years ago. We identified over 20 concrete acts a person could perform regarding a product, expressed in general terms. These were combinations of four acts (uses, gives as gift, seeks advice about, gives advice about) and five "other people", intended to represent the five types of interpersonal acts identified by Triandis' research: self (private), sweetheart, boss, employee, and acquaintance. About two dozen subjects judged their intent to perform each of the acts regarding stimulus products described by wide variations in four general factors: familiarity, expense, whether or not use or consumption would be publicly conspicuous, and whether or not the product was closely related to the person's self-image. Obviously, our motivation was not too dissimilar from Raju's. We never wrote up the results of the factor analysis performed on the behavioral intent reports, because of the small sample size and because the response task fatigued subjects and this may have influenced their responses somewhat. Our research question was whether the inherent nature of the act (using, giving, seeking information, giving information) or the nature of the people who are part of the act attracts people's attention. For what they are worth, our data showed that the inherent differences between the acts attracted little interest. The basis for discrimination was the degree of intimacy implied by the person(s) involved in the acts.

A second question of research strategy concerns the use of cross-subject correlations rather than within-subject correlations. With 20 stimulus products, multiple observations are available for each of Raju's subjects. Even allowing for a reduction due to "product irrelevance", it seems possible to test the covariation between limited sets of the variables on a within-subject basis. This would yield an individual-level test of a basically individual level theory, and could alleviate problems in interpreting cross-sectional correlations caused by individual differences in scale usage. Whether or not one should use within-subject or cross-sectional designs and data analysis methods is a complex issue but since a within-subject design was used for data collection, within-subject data analyses might be productive.

#### Harvey: Evaluative Conflict and Information Search in the Adoption Process

Our concern must be with the theoretical arguments Harvey develops, since an empirical test is not discussed. Harvey defines a construct called "evaluative conflict" in terms of the disparity between the values associated with a product's attributes, then uses Kelman and Baron's (1968) taxonomy of conflict resolution mechanisms to derive the propositions that (i) evaluating a high-risk innovation is conceived as a "single-goal" decision problem in which people are motivated to resolve the disparity between attributes by "confronting" it, and (ii) that this implies extensive search for external evidence from trusted sources.

As a prelude, one might note that Kelman and Baron's original propositions about conditional modes for inconsistency resolution have never been tested empirically, to the author's knowledge. Since the theorizing is interesting, one must assume that the lack of testing stems from difficulties with interpreting or operationalizing what K & B mean by "single goal" or "dual goal" situations or by "denial", "compartmentalization", "transcendence" etc. It is natural therefore to examine how Harvey deals with this translation.

Harvey confines his interest explicitly to high risk innovations, and asserts that the associated decision process is usually a "single-goal" process. This means, using Harvey's wording, that the discrepant elements (i.e., the disparate product characteristics) are tied to the same goal state. In a two goal (or "n"-goal) situation, the disparate product attributes would be linked to different goal states.

The author would restate what K & B had in mind this way: Assume there are two dimensions, A and B, being used for the evaluation. A is related to one type of goal or consequence (e.g., personal health), and B to a different type of consequence (e.g., aesthetic pleasure). One type of conflict that can arise is if one receives disparate evidence about the product's ability to cause "aesthetic pleasure"; some evidence suggests that it will and some that it won't. This may be termed "within-dimension uncertainty". Another type of conflict arises if one thinks the product is good on one dimension and bad on the other. One knows with great certainty that the product has a good and a bad side to it. We can call this "cross-attribute disparity".<sup>1</sup> "Within-dimension uncertainty" corresponds, in my thinking, to K & B's "single goal" conflict; "cross-attribute disparity" corresponds to a "multi-goal" conflict.

If the above discussion captures what K and B meant in terms of the structure of one's beliefs about a product's attributes, then Harvey's "evaluative conflict" construct seems to be defined as cross-attribute disparity. If so, his domain is by definition a two-goal (or n-goal) conflict situation, and this would lead him (at least, using K and B taxonomies) to view "maintenance/confrontation" strategies, not "reduction/confrontation" strategies as the preferred mechanisms for handling evaluative conflicts stirred by risky innovations. More discussion about the link between Harvey's evaluative conflict notion and K and B's single- and dual-goal notions seems desirable or, more directly, about the way people deal with cross-attribute disparity vs within-dimension uncertainty .

In highlighting external information search as a preferred strategy for conflict resolution, Harvey actually introduces a quite different resolution strategy into K and B's scheme. K and B's 12 modes all are intended to represent strategies of self-initiated cognitive adjustment (rethinking the problem), unaided by external data. The sole exception seems to be an "action change", but this does not require new data. If, as Harvey argues, a single-goal, central value situation is of interest, K and B propose as "reduction/confrontation" modes a change in one's global evaluation of the product, a change in one's own action toward the product, a change in one's norm, or an attempt to persuade one's source of information that he/she is wrong, or to change the product itself. None of these involves a search for more information. Nor do any of the strategies K and B suggest for dual-goal, central value situations (bolstering, differentiation, transcendence) involve a search for new external data.

So Harvey in fact uses K and B's structure up to a point, then reasons beyond the internal modes of resolution they deal with to modes involving external search. This seems like a potentially productive line of reasoning, but it requires more development. For example, many of the mechanisms K and B describe could, perhaps, include some external information search (source derogation, compartmentalization, attitude change, bolstering, for example), but these appear in every quadrant of their taxonomy. It seems useful to reason in as much detail about information search strategies as K and B did about cognitive readjustment strategies. A proposition that more external information will be sought from trustworthy sources seems too general. What types of information seeking activities do the respective situations K and B describe elicit? With what purpose, specifically? Why would one seek more data if one experienced cross-attribute disparity? To find another dimension to act as a tie-breaker? To make oneself more uncertain about one attribute? What if one's calculation of cross-attribute disparity already encompasses the major evaluative dimensions?

In summary, it seems quite useful to integrate theorizing about conditional information search strategies with K and B's interesting theorizing about self-initiated cognitive adjustment strategies, and Harvey points us in that direction. More detailed reasoning regarding the nuances of information search activities, perhaps in collaboration with Raju, should prove very enlightening. Then, the nontrivial problems of operationally distinguishing the situations and especially the varied types of information search activities can be confronted.

<sup>1</sup>These terms evolved from discussions earlier discussions with Fredrick Winter.

- Harvey, J.W., "Evaluative Conflict and Information Search in the Adoption Process", paper presented at the Association of Consumer Research Conference, Miami, 1978.
- Hendrick, C., Mills, J., and Kiesler, C.S., "Decision Time as a Function of the Number and Complexity of Equally Attractive Alternatives", Journal of Personality and Social Psychology, 8 (1968), 313-318.
- Jacoby, J., Speller, D., and Kohn C., "Brand Choice as a Function of Information Load: Replication and Extension", Journal of Consumer Research, 1 (1974), 33-42.
- Jacoby, J., Szybillo G., and Busato-Schach, "Information Acquisition in Brand Choice Situations", Journal of Consumer Research, 3 (1977), 209-216.
- Kelman, H.C. and Baron, R.M., "Determinants of Modes of Resolving Inconsistency Dilemmas: A Functional Analysis", in R. Abelson et al (Eds.), Theories of Cognitive Consistency: A Sourcebook. Chicago: Rand McNally, 1968.
- Kiesler, C.A., "Conflict and the Number of Choice Alternatives", Psychological Reports, 18 (1966), 603-610.
- Raju, P.S., "Stimulus-Response Variables in New Product Research", paper presented at the Association for Consumer Research Conference, Miami, 1978
- Reibstein, D.J., Youngblood, S.A. and Fromkin, H.L., "Number of Choices and Perceived Decision Freedom as a Determinant of Satisfaction and Consumer Behavior", Journal of Applied Psychology, 60 (1975), 434-437.
- Triandis, H.C., The Analysis of Subjective Culture. New York: Wiley, 1972.
- Van Raaij, F., Consumer Choice Behavior: An Information Processing Approach. PHD Dissertation, Tilburg (Hol.) University, 1977.
- Walton, J.R. and Berkowitz, E.N., "The Effects of Choice Complexity and Decision Freedom on Consumer Choice Behavior", paper presented at the Association for Consumer Research Conference, Miami, 1978.
- Wright, P., "Decision Times and Processes on Complex Decision Problems", Faculty Working Paper, Graduate School of Business, Stanford University, 1978.

Frederick E. May, University of Missouri-St. Louis<sup>1</sup>

Abstract

This discussion paper examines the controversy between learning theory and information processing theory in terms of their abilities to explain evoked set composition and formation. These two hypotheses appear to lead to contradictory expectations with respect to the composition and formation of evoked sets. The report examines existing data on evoked set formation in an attempt to reconcile these two approaches.

Introduction

This report is concerned with Howard's (1977) concept learning theory and McGuire's (1976) information processing theory in relation to consumer choice of evoked sets. For the purpose of this discussion the evoked set is "defined as those brands the buyer considers when he (or she) contemplates purchasing a unit of the product class." (Howard and Sheth, 1969; Glossary of Terms: Definitions of Central Concepts, p. 416).

The evoked set is important because it is a subset of existing brands in the market, and it is a subset of the brands of which the buyer is aware. The process of forming an evoked set from these broader sets is therefore of interest to both consumer researchers and managers.

McGuire postulates that the choice of the evoked set is an intermediary decision toward the brand choice and that the process of identification is different from the process of choosing a brand within the set.

Ideally, one would like to identify by name the brands that the consumer selects for the evoked set. However, a more practical approach is to identify brands in the evoked set by the consumer's prior experience with them. Thus the dependent variable becomes the following typology of evoked sets: those evoked sets composed solely of untried brands, those composed of some tried and some untried brands, and finally, those composed solely of tried brands.<sup>3</sup>

The composition of the evoked set as a dependent variable permits a comparison of the learning and information processing approaches to evoked set formation. Learning theory predicts that all consumers progress from untried sets through mixed sets to tried sets as

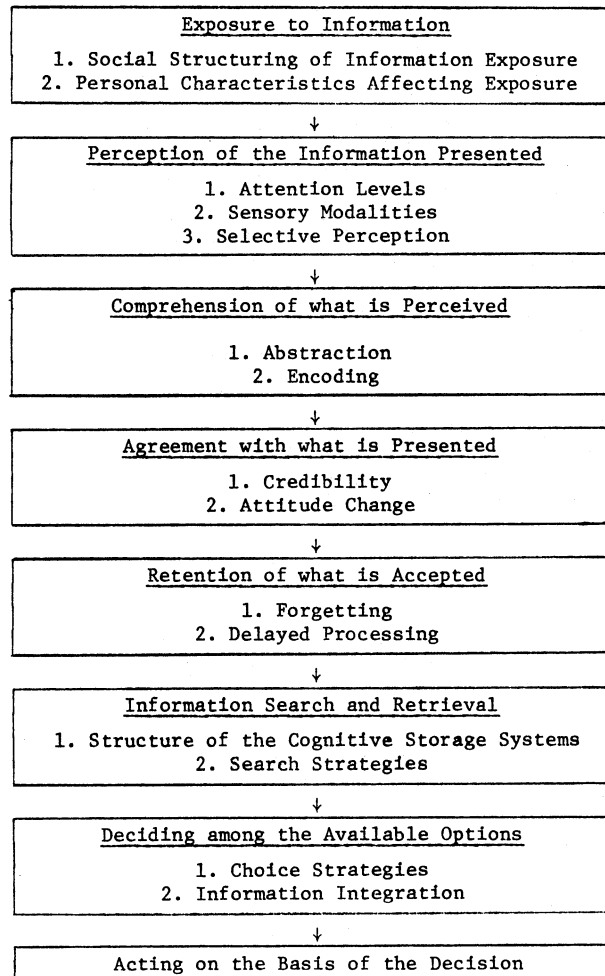
they acquire experience with the product. Information processing theory states that the composition of the consumer's evoked set depends on the social and personal characteristics of the consumer and the decision rules used to identify the evoked set.

McGuire's Information Processing Model

McGuire's (1976) model of directive aspects in personality structure is shown in Figure I. There are eight stages in the information processing sequence. In each stage McGuire postulates two or three constructs. McGuire's entire paradigm is too massive to deal with in this report. This report focuses on two steps---exposure to information, and information search and retrieval.

FIGURE I

INFORMATION PROCESSING IN CONSUMER DECISION MAKING



There are two constructs presented in search and retrieval. The available options (evoked set) are identified or attained by search strategies. The structure of

<sup>1</sup>Thanks to Johann Arndt, and Ronald P. May for their helpful suggestions.

<sup>2</sup>McGuire defines evoked set as the available options that are generated as a result of the search strategy in the search and retrieval step in the information processing sequence. It is assumed that the options correspond to the brands considered by the buyer.

<sup>3</sup>Howard and Sheth (1969) develop a typology of the evoked set based on the consumer's familiarity with the brands considered in order to predict the amount and duration of search that would follow the formation of the evoked set. Here, one is interested in using a similar typology (the operational equivalent of the familiarity dimension) in order to analyze the factors leading to the formation of the evoked set.



the cognitive storage systems refers to tree diagrams, matrices, and push down lists.

McGuire presents two constructs in the exposure to information stage. For the social structuring of information exposure the adoption and diffusion of innovation literature is relevant. This is recently reviewed by Engel, Blackwell, and Kollat (1978). The cognitive style literature reviewed by Pinson (1975) provides a valuable discussion of the personal characteristics affecting exposure to information.

#### Social and Personal Characteristics

The social characteristics of the consumer which direct the consumer's exposure to information here refer to the consumer's position in the social structure. An operational measure of social position is occupational prestige. An analysis of variance shows a significant relationship between evoked set composition and occupational prestige in a study of new automobile buying (May, Homans, and Maddox, 1977b). The untried evoked sets are correlated with the highest occupational prestige and the tried sets with the lowest. Income was found to be a weak predictor of evoked set composition.

Personal characteristics here refer to information processing complexity of which one important indicator is education. Pinson (1975) describes the three basic aspects of information processing complexity:

"Differentiation', 'Discrimination', and 'Integration'. Broadly defined, 'differentiation' refers to the number of dimensions used by an individual in processing information. 'Discrimination' refers to the number of separated conceptual categories on a dimension. Finally, 'integration' refers to the degree of interrelatedness of elements within a particular cognitive domain."

Although studies have not linked education directly to information processing complexity, education is a frequent and strong correlate in studies of the amount of information search (Newman, 1977). Amount of information search has in turn been shown to be correlated with information processing complexity in studies reviewed by Pinson (1975).

Education and occupational prestige correlates with evoked set composition and size in a canonical correlation analysis of new car buyers (Homans, Maddox, and May, 1977). Evoked set size, which tends to be correlated with composition, has been shown to be related to education in studies of Norwegian and United States car buyers (Maddox, Grønhaug, Homans, and May, 1978). Brand loyalty is inversely related to education across grocery product categories (Frank, Douglas, and Polli, 1968). These data, however, speak only indirectly to the composition of the evoked set.

#### Consumer Product Specifications

Product specifications are a third major factor that determines the composition of the evoked set. Product specifications are the appropriate decision rules for forming the evoked set. In the study of industrial buyer behavior the use of specifications is well known. In fact, the terminology used here comes from the industrial buying literature (Lamar and Dobler, 1971). However, the notion that the consumer constructs specifications as a search strategy is new to the consumer behavior literature (Bettman and Zins, 1977).

May, Homans, and Maddox (1978) find four types of product specifications: 1) single, and 2) multiple brand category specifications, 3) attribute specifications, and 4) performance specifications. New car buyers were

asked: "When you first decided to buy what were your specific ideas about the type of car you wanted?"

The answers were classified in the following manner: the respondent mentioned the name of a single brand (single brand category specifications); the respondent mentioned the name of more than one brand (multiple brand specifications); the respondent mentioned attributes of the car wanted (attribute specification). For example, the respondent may have mentioned that he wanted a small car, a heavy car, a four door sedan.

If the respondent mentioned a function or activity for which the car is to be used and related it to a necessary attribute he was placed in the performance specification category. For example, the respondent may have expressed the desire for an economical car for city driving, or a station wagon for a big family and hauling things, or a heavy car for highway driving (May and Homans, 1977a).

The product specifications range over a dimension of complexity from the simplest (single brand) to the most complex (performance) specifications. The choice of product specification has been shown to be a function of personal and social characteristics, of which education correlates most highly with the complexity of specifications (May, Homans, and Maddox, 1978).

The different specifications are likely to produce different types of evoked sets. To illustrate: when brand specifications are used the most frequent outcome is a tried evoked set, and when attribute specifications are used, the most frequent outcome is a mixed evoked set. Finally, when a performance specification is used the most frequent outcome is an untried set.

The product specifications appear to correspond to three types of memory structures upon which the consumer draws to construct the evoked set. These structures are utilized in the search and retrieval step of McGuire's information processing model.

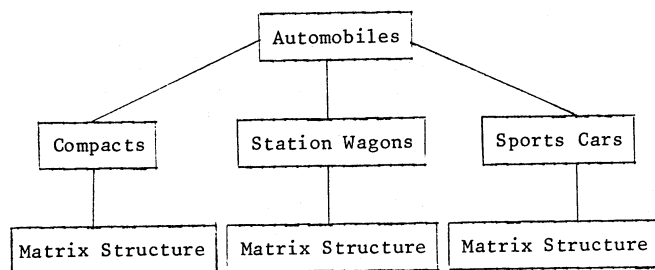
#### Memory Structures

McGuire's hypothesized memory structures provide a conceptual explanation of the link between product specifications and evoked set composition. The three memory structures are shown in Figure II. The hierarchical network product concept (tree-diagram structure) encompasses the entire product class which is usually labeled by nouns and sub-product classes that are labeled by using adjectives. For example, the superordinate product class concept of automobiles is sub-ordinated by such classes as compact cars, station wagons, sports cars, etc. Each of the sub-product classes is usually designated by an adjective and has a distinguishable function. For example, compacts are for economy, station wagons are for large families, and sports cars are for fun and pleasure.

Matrix product concept structures refer to brands and their attributes. Such structures are likely to be used for one product or sub-product class. Finally, push down lists (or preference lists) are structures which contain a list of preferred or acceptable brands. The list is a kind of collapsed matrix. The consumer buys the first preference if it is available at the right price. If not available, the second preference is selected, and so forth.

The particular memory structure in which the consumer stores relevant information about the product influences the type of information which the consumer seeks about the product in the process of forming the evoked set. To illustrate, the most appropriate memory structure from which to draw performance specifications is a network structure because it defines the domain (product

FIGURE II  
NETWORK PRODUCT CONCEPT STRUCTURES



Matrix Brand X Attribute Structures

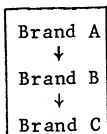
	Brands				
Attributes	A	B	C	D	E
X					
Y					
Z					

Push Down List

First Preference if available

Second Preference if first is not available

Third Preference if special deal is offered



class) in functional terms. Following logically, the most appropriate memory structure for the attribute specifications is the matrix. The matrix structure can be a sub-structure in the network structure. Finally, the push down list is the most appropriate memory structure for a single or multiple brand product specification.

#### Need For a Dynamic Theory

One problem with the information processing paradigm is that it assumes a static consumer who will, once his social and personal characteristics are determined, always follow the same pattern. The highly educated consumer, for example, can be expected to always construct network memory structures, performance specifications, and untried evoked sets. How does the highly educated consumer ever come to examine tried sets and how does the poorly educated consumer ever come to examine untried sets? Stated differently, how does the poorly educated consumer adopt a new brand or product? To answer this question we need a theory of dynamics or learning.

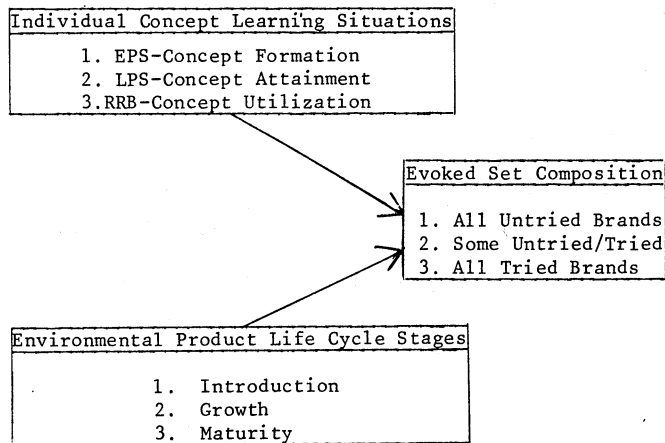
#### Howard's Theory of Concept Learning

Howard (1977) distinguishes three concept learning phases in the adoption process. (See Figure III). In the first phase the consumer forms the concept of a new product (or sub-product) by comprehending its distinctive functions and attributes. Howard calls this phase concept formation. Concept formation is here defined as the learning process that results in untried sets.

In the second phase of new product adoption the consumer extends the product concept by comparing untried brands to previously tried brands. Howard calls this process concept attainment. However, psychologists generally do not discriminate between concept formation and attainment. Therefore, the term concept extension

is more appropriate for this phase of the adoption process. Concept extension results in mixed evoked sets.

FIGURE III  
MODEL OF EVOKED SET COMPOSITION



In the third phase of the adoption process the consumer has completely learned the product concept, and begins to utilize it. In this phase the consumer chooses only tried brands in the evoked set. Howard calls this phase concept utilization.

To predict the frequency distribution of concept learning situations, Howard introduces the product life cycle. In the introduction stage most consumers are in the concept formation phase and, therefore, untried evoked sets are most common. In the growth stage most consumers are in the concept extension phase and mixed sets of some tried and untried brands are most frequent. In the maturity stage most consumers are in the concept utilization stage, and tried sets are most common.

There is some evidence that consumers proceed through the three adoption phases and form evoked sets according to the theory. For example, the mean age and the mean number of previous purchases for the different types of evoked sets appear to vary as predicted (May, Homans, and Maddox, 1977b). However, variance analyses indicates that the age and purchase distributions do not differ significantly.

One can conclude that age and the number of prior purchases are confounded by education and occupational prestige factors. That is, given the information processing complexity of consumers the learning hypothesis appears to apply.

Howard's product life cycle hypothesis predicts the frequency distribution of evoked sets correctly. New car buyers considered tried evoked sets most frequently, and untried sets least frequently (May, Homans, and Maddox, 1977c). The finding is somewhat surprising in view of Howard's argument that durable goods buyers are unlikely to reach the concept utilization phase even in the maturity stage of the product life cycle.

#### Discussion

Table I summarizes typical consumer behavior according to the two theoretical approaches. Howard's three stage process corresponds in certain respects to the three levels of information processing complexity. The

TABLE 1

## TYPICAL CONSUMER BEHAVIOR IN INFORMATION PROCESSING STEPS AND CONCEPT LEARNING SITUATIONS

Information Processing Steps	Concept Formation	Concept Extension	Concept Utilization
I. EXPOSURE			
A. Social Structure	High social status, early family and product life cycle	High S.E.S; middle l.c.	Low S.E.S; late family and product life cycle
Education	High	Intermediate	Low
Occupational Prestige	High	Intermediate	Low
Age	Young	Middle	Old
Income	Middle	High	Low
B. Cognitive Personality Structure	Abstract; Complex; Integrative and Analytical	Intermediate Differentiative Analytical	Concrete; Simple; Discriminant; Global Categorical
C. Information Sources User	Acts like Innovator	Acts like Early Adopter	Acts like Late Adopter
II. SEARCH AND RETRIEVAL			
A. Retrieves	Superordinate product concept-Hierarchical network	Subordinate product concept-Matrix	Preference ranking of acceptable brands-Push down list
B. Constructs	Performance specifications	Attribute specifications	Brand category specifications
C. Identifies Evoked set	Evoked set of untried brands that fit specs.	Untried and tried brands	Tried brands only

concept formation phase requires the highest level of information processing complexity, and the utilization phase requires the simplest level of information processing.

Howard's theory assumes that the accumulation of product knowledge and experience changes the consumer's information processing behavior. Learning simplifies the information processing task.

In contrast to learning theory, information processing theory assumes that the consumer starts at a given level of complexity and remains there. The theory implies that the type of product specifications used by consumers do not change as the consumer accumulates purchase experience. So far unpublished data support the latter hypothesis. The proportion of brand specifiers remains constant, regardless of the number of prior new car purchases.

The implication for the frequency distribution of evoked sets is as follows. Information processing theory predicts that the frequency distribution of evoked sets remains constant as long as the frequency distribution of the complexity of the buyers of the product remains constant. As the product diffuses through the social structure one would expect the frequency of evoked set compositions to change as less complex buyers adopt the product.

Learning theory predicts that the frequency distribution of evoked sets remains constant as long as the familiarity of the buyers with the product remains constant. As the product progresses through the life cycle one would expect the distribution of evoked sets to change with

increasing buyer familiarity with the product.

One may conclude that both theories are necessary to predict and explain evoked set composition and formation.

## References

- Bettman, J. R. and Zins, M. A. (1977), "Constructive Processes in Consumer Choice," Journal of Consumer Research, 4, 75-85.
- Engel, J. F., Blackwell, R. D., and Kollat, D. T. (1978) Consumer Behavior, Hinsdale, IL: Dryden Press.
- Frank, R. E., Douglas, S. P. and Polli, R. E. (1968) "Household Correlates of 'Brand Loyalty' for Grocery Products," Journal of Business, 41, 237-245.
- Homans, R. E., Maddox, R. N., and May, F. E. (1977) "Correlates of the Level of Decision Processing for Automobiles," Proceedings of the American Institute of Decision Sciences.
- Howard, J. A. (1977), Consumer Behavior: Application of Theory, New York: McGraw-Hill.
- \_\_\_\_\_, and Sheth (1969) The Theory of Buyer Behavior, New York: Wiley & Sons, Inc.
- Lamar, L. Jr. and Dobler, D. W. (1971) Purchasing and Materials Management, New York: McGraw-Hill.
- McGuire, J. (1976) "Some Internal Psychological Factors Influencing Consumer Choice," Journal of Consumer Research, 2, 302-319.
- Maddox, R. E., Grønhaug, K., Homans, R. E. and May, F. E. (1978) "Correlates of Information Gathering and Evoked Set Size for New Automobile Purchasers in Norway and the U.S." in Advances in Consumer Research, Volume V, Hunt, K. ed., Ann Arbor: Association for Consumer Research.
- May, F. E., and Homans, R. E. (1977a) "Evoked Set Size and the Level of Information Processing in Product Comprehension and Choice Criteria," in Advances in Consumer Research, Vol. 4, ed. William D. Perreault, Jr., Ann Arbor: Association for Consumer Research.
- \_\_\_\_\_, Homans, R. E., Maddox, R. N. (1977b) "Dimensions of Problem Solving Behavior in New Automobile Purchases," in Proceedings of 4th International Research Seminar in Marketing, Senanque Abbey, France: Institut D'Administration des Enterprises.
- \_\_\_\_\_, Homans, R. E., and Maddox, R. N. (1977c) "Evoked Sets and Memory Systems" in Contemporary Marketing Thought 1977 Educators Proceedings, Greenberg and Bellenger, eds. Chicago: American Marketing Association.
- \_\_\_\_\_, Homans, R. E., and Maddox, R. N. (1978) "Evaluative Criteria and Cognitive Style," Proceedings of the Southwestern Marketing Association, Dallas, Texas.
- Newman, J. W. (1977) "Consumer External Search: Amount and Determinants," in Consumer and Industrial Buying Behavior, eds. Woodside, A. G.; Sheth, J. N.; Bennett, P. D., Amsterdam: Elsevier North-Holland, Inc.
- Pinson, C. (1975) "Consumer Cognitive Styles: Review and Implications for Marketers," Research Paper No. 160 INSEAD The European Institute of Business Administration.

AN INFORMATION PROCESSING APPROACH TO  
EVOKED SET FORMATION

Thomas L. Parkinson, The Pennsylvania State University  
Michael Reilly, The Pennsylvania State University

Abstract

This exploratory study investigates the dynamics of evoked set formation from an information processing perspective. Subjects were asked to identify actual evoked sets and to provide the necessary attribute rating data for simulating an evoked set decision utilizing five information processing strategies. When the actual and simulated evoked sets were compared, the best "fits" resulted from application of the unweighted linear compensatory and lexicographic strategies.

Problem

With the proliferation of brands that has occurred in the marketplace, it is well known and widely accepted that a consumer only considers a subset of the available brands when making a purchase decision. Howard and Sheth (1969) termed this subset of "acceptable" brands the evoked set, and this concept has received widespread acceptance in marketing in spite of its lack of empirical validation. The research which has been reported concerning the evoked set has been static in nature focusing primarily on the situational and individual factors that influence the evoked set size of an individual in a given purchase situation.

Campbell (1969), Ostland (1974) and Grønhaug (1973/74) investigated the relationship of evoked set size and a large number of standard marketing variables, demographic characteristics, and selected individual difference variables. In general, these studies failed to produce significant results, and in most cases only a small percentage of the variance in evoked set size was explained. Only brand loyalty and level of education were found to be related to evoked set size. Using social judgment theory as a basis, Jarvis and Wilcox (1973) investigated the relationship between perceived importance of the product class and evoked set size; and, as hypothesized, demonstrated a positive relationship between the two. These findings coupled with the work of May and Homans (1977) which demonstrated a relationship between the abstractness of an individual's information processing (cognitive style) and evoked set size support the proposition that the size of the evoked set is at least in part an individual variable, related to cognitive style.

The evidence that the size of the evoked set is related positively to education level and degree of abstractness in product perception; together with other studies which have suggested the existence of multiple subsets in addition to the evoked set (Jacoby, 1973; Narayana and Markim, 1975) suggests that the formation of evoked sets is a cognitive process amenable to study from an information processing perspective. Recently, Belonax and Mittelstaedt (1978) suggested such an approach and postulated a two-stage model with evoked set formation as a prerequisite for further brand evaluation. However, they too utilized evoked set size as the dependent variable, and found that set size was inversely related to the number of choice criteria and information variability. Thus in spite of the significance of the evoked set to marketing practitioners, to date there has been no attempt to study the dynamic process of evoked set formation.

Information Processing

This lack of a dynamic approach can, perhaps be traced partly to a lack of a sound theoretical base for such an inquiry. However, recent developments in the field of consumer information processing would appear to provide the tools for this research. Study of human information processing originated in the study of human clinical judgment (Dawes, 1971, 1972; Einhorn, 1970, 1971; Goldberg, 1968, 1970, 1971; Hoffman, 1960, 1968), and in the attempts to program computers to simulate human problem solving (Newell, Shaw and Simon, 1958; Newell and Simon, 1972; Simon and Newell, 1970). For recent reviews of this research and its applicability to marketing problems see Bither and Ungson (1975), and Jacoby and Chestnut (1977).

One of the major contributions of this research has been the isolation of a number of algorithmic judgment strategies. These strategies notably weighted and unweighted linear compensatory, conjunctive, disjunctive, and lexicographic or elimination by aspects have been hypothesized to represent consumer judgment strategies during purchase behavior, and several methods for measuring them have been developed (Jacoby, Speller, and Kohn-Berning, 1974, 1974b; Einhorn, 1970; Goldberg, 1971; Bettman, 1974; Wright, 1975; Reilly and Holman, 1977).

If one views the formation of an evoked set by a consumer when faced with a multitude of brands as an attempt to simplify the decision making environment (Miller, 1956; Wright, 1974, 1975), then the consumer may use the previously discussed strategies to accomplish this. By only considering a subset of the available brands in a purchase situation, a consumer would only be required to process a fraction of the available product-related information. Thus, the existence of an evoked set for a product category would seem to imply the operation of at least a two-stage decision strategy. In the first stage the individual decides which brands to consider (i.e., forms the evoked set) by the application of a processing strategy. Then, when a purchase situation arises the consumer applies another strategy, or possibly reapplies the same strategy, to the elements of the evoked set to make a decision.

In an actual temporal sense what probably happens is that the consumer considers one brand at a time and makes a binary decision as to whether or not the brand is worthy of future consideration. If the brand is deemed suitable it enters the evoked set. If not, it is rejected. This binary decision is made as soon as the consumer feels that he has enough information to evaluate the brand. The purposes of this research is to investigate which processing strategy is most frequently used by consumer in formulating their evoked set for a given product class. Specifically, the following five strategies will be evaluated for their ability to approximate actual evoked set decisions.

1. Unweighted Linear Compensatory: When utilizing this strategy the individual is seen as summing the attribute ratings for each available brand and then including all brands whose final ratings exceed a certain cutoff in the evoked set.

2. Weighted Linear Compensatory: When utilizing this strategy the individual is seen as summing the attribute

ratings, weighted by their importance, for each available brand and then including all brands whose final ratings exceed a certain cutoff in the evoked set.

3. Conjunctive: When utilizing this strategy the individual is seen as including in the evoked set only those brands whose attribute ratings exceed predetermined cutoffs on all attributes.

4. Disjunctive: When utilizing the strategy the individual is seen as including in the evoked set all brands which exceed a predetermined cutoff on at least one attribute.

5. Lexicographic: When utilizing this strategy the individual is seen as rank ordering the attributes of the brands on an importance dimension, and then ranking all available brands based on their rating on the most important attribute. Ties are seen as being broken by using the second most important attribute and so forth. Then only brands which exceed a certain cutoff are included in the evoked set.

#### Methodology

In order to investigate the possible application of various information processing strategies to the formation of an evoked set, a written questionnaire was designed and administered to 90 undergraduate students enrolled in three sections of the University's consumer behavior course during the spring term. The instrument first asked the subjects to identify their evoked sets for two product categories and then to provide the additional data necessary for an individual actuarial analysis of the ability of the postulated strategies to correctly duplicate the actual evoked sets of the respondents. The two products selected for inclusion in the design were toothpaste and underarm deodorant. These were selected based on their high level of use and familiarity among students. However, each subject was asked about only one product category due to the length of the questionnaire.

The actual evoked set for each product category was established by presenting the subjects with a list of available brands (8 brands of toothpaste; 12 brands of deodorant), and asking them to indicate those brands with which they were familiar (the awareness set). They were asked to indicate which brands in the awareness set they would consider buying if faced with an immediate purchase decision. This second measure was identified as the actual evoked set for the product category.

Once the actual evoked set was identified, the subjects were asked to evaluate each brand in the awareness set on a number of attributes. These attributes were selected based on previous research (Haley, 1968; Ryan, 1978) and an analysis of current promotional literature. The following attributes were included:

<u>TOOTHPASTE</u>	<u>DEODORANT</u>
Taste	Smell
Decay Prevention	Ease of Application
Bad Breath Prevention	Effectiveness
Abrasiveness*	Likelihood of Staining Clothes*
Mouth Refreshment	Length of Protection
Teeth Whitening Ability	Economy in Use
Price*	Skin Irritation*
	Price*

Each of the attributes was measured utilizing a seven point semantic scale with those attributes marked with an \* being reverse scored due to their perceived negative influence on purchase behavior.

When all of the familiar brands were evaluated, then the subjects were asked to rate the importance of the attributes themselves with respect to their use in evaluating brands in the product category under consideration. Again, a seven point semantic scale was used. Following the attribute importance determination, the conjunctive and disjunctive cutoff points for each attribute were measured. For this purpose the following question formats were utilized:

#### Conjunctive Cutoff.

If I found that the Effectiveness of a brand was rated below (Mark the appropriate point)

Effectiveness

Very \_\_\_\_\_: \_\_\_\_\_: \_\_\_\_\_: \_\_\_\_\_: \_\_\_\_\_: \_\_\_\_\_: \_\_\_\_\_ Very  
High \_\_\_\_\_: \_\_\_\_\_: \_\_\_\_\_: \_\_\_\_\_: \_\_\_\_\_: \_\_\_\_\_: \_\_\_\_\_ Low

I would not consider purchasing that brand regardless of how the brand was rated on other attributes.

#### Disjunctive Cutoff.

If I found that a brand was rated at or above this point with respect to the attribute Effectiveness, I would consider purchasing it regardless of how the brand was rated on all other attributes. (Mark the appropriate point on the scale below)

Effectiveness

Very \_\_\_\_\_: \_\_\_\_\_: \_\_\_\_\_: \_\_\_\_\_: \_\_\_\_\_: \_\_\_\_\_: \_\_\_\_\_ Very  
High \_\_\_\_\_: \_\_\_\_\_: \_\_\_\_\_: \_\_\_\_\_: \_\_\_\_\_: \_\_\_\_\_: \_\_\_\_\_ Low

OR  
The attribute effectiveness alone (regardless of how rated) is not a sufficient condition for considering a brand for purchase \_\_\_\_\_ (check if applicable).

Following completion of the final part of the study, the students were introduced to the theoretical basis for the research and encouraged to ask questions about the study.

#### Analysis

In order to determine the fit of each information processing strategy, an evoked set decision was simulated utilizing each of the five models and the results compared with the actual evoked sets identified at the start of the questionnaire.

A total of 90 subjects took part in the study divided evenly between the two products. Two questionnaires both dealing with toothpaste were judged unusable due to missing responses and were discarded. Furthermore only subjects who rated attributes of three or more brands, and identified at least one brand in the evoked set and at least one brand in the non-evoked set were included in the analysis. This latter criterion was made necessary by the inclusion of the compensatory and lexicographic models which do not produce discrete evoked set decisions. This further elimination of subjects from consideration reduced the number of usable responses to 30 to 27 for toothpaste and deodorant, respectively.

The stimulated decision making procedure produced two types of output. In the cases of two linear compensatory and the lexicographic strategies a ranking of the brands was achieved. Brands ranked higher than the size of the actual evoked set were identified as members of the

stimulated evoked set. When the conjunctive and disjunctive strategies were applied to the data the result was not a ranking, but a binary decision directly identifying membership or non-membership in the evoked set.

Once the stimulated evoked and non-evoked sets were specified, the results using each processing strategy were compared to the actual evoked set specified at the outset by the subject.

### Results

Two measures of "success" were utilized to evaluate the performance of the five information processing strategies and the results are reported in Table 1A and 1B. In Table 1A a "best fit" criterion was applied. The similarity of the simulated and actual evoked sets was evaluated based on the percentage of correct decisions. Those strategies which produced the highest percentages (ties included) of correct answers were credited with a "successful" result. In Table 1B the more demanding "perfect" criterion was used. Only simulated evoked sets which matched actual evoked sets exactly were counted as successes. This more stringent criterion produced absolutely lower success rates, but the relative performance of the different processing strategies remained the same.

### Discussion

Any discussion of the results of the current study must first recognize the exploratory nature of the research. Its major contribution lies in the conceptualization of the formation of the evoked set as dynamic, and the adoption of information processing theory as a basis for studying this process. Thus, the critical aspect of the paper is theoretical rather than empirical.

### Limitations

There are three limitations of this initial study which may be operating to bias the outcomes. These include the conceptual base, the methodology and the data analysis procedure.

The shortcomings of information processing theory as a conceptual base for studying a decision whether or not to include a brand in the evoked set, are essentially the same as the limitations of information processing theory in the study of other decision processes. Specifically, there is some question concerning the use of information processing strategies as the basic unit of analysis. The idea that individuals in a decision situation use these relatively abstract and complex strategies has never received extensive empirical validation. In reality the decision process may be much more idiosyncratic, and therefore, more difficult to represent algorithmically. On the other hand these processing strategies seem to have a fairly high potential for simplifying inquiry into the decision process due to their ability to approximate a variety of different decisions. Accordingly, the use of these strategies in the current research should not be too debilitating in light of the exploratory nature of the research.

The other limitations of the current study are in the area of data collection and analysis. These shortcomings are, of course, far more significant when interpreting the results, than to the theoretical basis. The major methodological problem arose while attempting to measure the conjunctive and disjunctive cutoffs. These measures appeared to be fairly difficult for the subjects to complete. This could, however, be due in part to the inappropriateness of these strategies in making an evoked set decision, and therefore, may not be a pure methodological problem. In either case, any response errors

TABLE 1A

Percentage of Successful Evoked Set Determinations:  
Best Criterion<sup>1</sup>

Product	n	Unweighted Linear Compensatory	Weighted Linear Compensatory	Conjunctive	Disjunctive	Lexicographic
Toothpaste	30	73(2)	50(3)	7	20	80(1)
Deodorant	27	70(1)	41(3)	4	15	59(2)

<sup>1</sup>Successful determination is defined as resulting in the highest percentage of matches between the actual and simulated evoked sets (ties included).

TABLE 1B

Percentage of Successful Evoked Set Determinations:  
Perfect Criterion<sup>2</sup>

Product	n	Unweighted Linear Compensatory	Weighted Linear Compensatory	Conjunctive	Disjunctive	Lexicographic
Toothpaste	30	43(2)	27(3)	0	0	50(1)
Deodorant	7	33(1)	19(3)	0	4	26(2)

<sup>2</sup>Successful determination is defined as resulting in a 100% match between the actual and simulated evoked sets.

of this type are probably magnified during the simulation procedure.

In addition, some problems may have arisen due to the fact that a significant number of subjects rated a very limited number of brands. Hopefully, this problem was reduced by including only data from subjects who rated at least three brands. However, when the number of brands rated was small, determination of the best fit between the simulated and actual evoked set was more difficult. Accordingly, care must be exercised in the interpretation of the results.

#### Findings

It is fairly obvious that the best fitting strategies in the current study were the lexicographic and the unweighted linear compensatory, regardless of which criteria for evaluation was used. Such a result is rather surprising in light of the differences between these two strategies. The lexicographic strategy places supreme weight on the most important attribute and other attributes are used only to break ties. The unweighted linear compensatory strategy on the other hand weights all attributes equally. It seems unusual that these two strategies consistently fit the data best. This result suggests two possible interpretations. One possible explanation for such a result could be that there were two segments of market, differing perhaps on level of involvement, who consistently used different strategies. For one group all of the attributes are equally important, while for the other group the most important dominated the decision.

A second possibility is that the result may be an artifact of the experimental procedure. The tendency of subjects to rate an object on a number of attributes in a manner that is congruent with their overall impression of the object has long been recognized. This tendency, termed halo effects has a long history in both psychological and marketing research (Symond, 1925; Thorndike, 1920; Bingham, 1939; Beckwith and Lehman, 1975, 1976; Johansson, MacLachlin and Yalch, 1976). For recent reviews see Beckwith, Kassarian and Lehman (1978) and Huber and James (1978). In the current study, it is a defensible position that the success of the unweighted linear compensatory model is due to the subjects' attribute ratings being effected by their previous evaluation of the brands. Clearly, additional research is needed to clarify this issue. One possibility is an experiment involving the introduction of a totally new brand which there is no prior evaluation. The evoked set formation process could then be analyzed without contamination by halo effects.

A second notable outcome is the failure of the weighted linear compensatory, the conjunctive and the disjunctive strategies to accurately simulate the contents of the evoked set. This is particularly surprising for the conjunctive and disjunctive strategies which, by their very nature, are adapted to making a binary decision on the acceptability of an alternative. However, due to the aforementioned limitations of the study it would be hasty to conclude at this point that these strategies are not representative of the decision process by which a consumer generates an evoked set. Again more research is needed before a definitive statement can be made.

#### Future Research

Assuming that the methodological and theoretical difficulties which hindered the current research can be overcome, and we think that they can; the area of evoked set research is one of great theoretical and practical value. The use of an information processing perspective enables the consideration of a large number of factors which may possibly influence the process of evoked set formation. However, more research is needed to investigate among other things:

1. The individual and situational factors which determine the choice of a processing strategy in the evoked set formation.
2. How the way in which information about a new brand is presented influences the outcome of the evoked set decision, for the particular brand.
3. The way in which a consumer decides to reject a familiar brand currently in the evoked set when confronted with an attractive new brand. This issue is important in light of past research demonstrating the relatively constant size of evoked sets.

#### References

- Neil E. Beckwith, Harold H. Kassarian, and Donald R. Lehmann, "Halo Effects in Marketing Research: Review and Prognosis," in H. Keith Hunt, (Ed.), Advances in Consumer Research, V (1978), 465-467.
- and Donald R. Lehmann, "The Importance of Halo Effects In Multi-Attribute Attitude Models," Journal of Marketing Research, 12 (August, 1975), 265-275.
- Joseph A. Belonax and Robert A. Mittelstaedt, "Evoked Set Size as a Function of Number of Choice Criteria and Information Variability," in H. Keith Hunt (Ed.), Advances in Consumer Research, V (1978), 48-51.
- James Bettman, "Toward A Statistics for Consumer Decision Net Models," Journal of Consumer Research, 1 (June, 1974), 71-80.
- W. V. Bingham, "Halo, Involved and Valid," Journal of Applied Psychology, 23 (1939), 221-229.
- Stewart W. Bither and Gerardo Ungson, "Consumer Information Processing Research, An Evaluative Review," Working Series in Marketing Research, No. 28, (April, 1975), The Pennsylvania State University.
- Brian M. Campbell, "The Existence of Evoked Set and Determinants of Its Magnitude in Brand Choice," Unpublished doctoral dissertation, Columbia University, 1969.
- Robyn M. Dawes, "A Case Study of Graduate Admissions: Application of Three Principles of Human Decision Making," American Psychologist, 26 (February, 1971), 180-188.
- , "Slitting the Decision Maker's Throat With Occom's Razor: The Superiority of Random Linear Models to Real Judges." Paper delivered at Seminar on Multiple Criteria Decision Making, Columbia, S. C., 1972.
- Hillel Einhorn, "The Use of Nonlinear, Noncompensatory Models in Decision-Making," Psychological Bulletin, 73 (March, 1970), 221-230.
- , "The Use of Nonlinear, Noncompensatory Models As A Function of Task and Amount of Information," Organizational Behavior and Human Performance, 6 (January, 1971), 1-17.
- Lewis R. Goldberg, "Simple Models or Simple Processes? Some Research On Clinical Judgments," American Psychologist, 23 (July, 1968), 483-496.
- , "Man vs. Model of Man: A Rationale Plus Some Evidence for A Method of Improving On Clinical Inferences," Psychological Bulletin, 73 (June, 1970), 422-432.
- , "Five Models of Clinical Judgment: An Empirical Comparison Between Linear and Non-Linear Representations of the Human Inference Process," Organizational Behavior: Human Performance, 6 (June, 1971), 458-479.



- Kjelf Grønhaug, "Some Factors Influencing the Size of the Buyers Evoked Set," European Journal of Marketing, 7 (Winter, 1973/74), 232-241.
- Russell I. Haley, "Benefit Segmentation: A Decision-Oriented Research Tool," Journal of Marketing, 32 (July, 1968), 30-35.
- Paul Hoffman, "The Paramorphis Representation of Clinical Judgment," Psychological Bulletin, 57 (March, 1960), 116-131.
- , "Cue Consistency and Configurality In Human Judgment," in B. Klermmutz, (Ed.), Formal Representation of Human Judgment, (New York: Wiley, 1968).
- John A. Howard and Jagdish N. Sheth, The Theory of Buyer Behavior, (New York: Wiley, 1969).
- Joel Huber and William James, "A Measure of Halo," in H. Keith Hunt, (Ed.), Advances in Consumer Research, V (1978), 474-476.
- Jacob Jacoby, "A Model of Multi-Brand Loyalty," Journal of Advertising Research, 11, 1971, 25-30.
- and Robert W. Chestnut, "Consumer Information Processing: Emerging Theory and Findings," in Woodside et al., (Eds.), Consumer and Industrial Buying Behavior, New York: Elsevier North-Holland, Inc., 1977, 119-134.
- , Donald E. Speller and Carol Kohn-Berning, "Brand Choice Behavior As A Function of Information Load: Replication and Extension," Journal of Consumer Research, 1 (June, 1974), 33-42.
- , Donald E. Speller and Carol Kohn-Berning, "Brand Choice Behavior As A Function of Information Load," Journal of Marketing Research, 1974b, 11 (February, 1974), 63-69.
- 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, 236-240.
- Johnny K. Johansson, D. J. MacLachlan and R. F. Valch, "Halo Effect in Multi-Attribute Models: Some Unresolved Issues," Journal of Marketing Research, 8 (February, 1976), 414-417.
- 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, 4, 1977, 172-175.
- George A. Miller, "The Magical Number Seven, Plus or Minus Two: Some Limits On Our Capacity for Processing Information," The Psychological Review, 63 (1956), 81-97.
- Chem L. Narayana and Rom J. Markin, "Consumer Behavior and Product Performance: An Alternative Conceptualization," Journal of Marketing, 39 (October, 1975), 1-6.
- Allen Newell and Herbert Simon, "Elements of A Theory of Human Problem Solving," Psychological Review, 65 (May, 1968), 151-166.
- and Herbert Simon, Human Problem Solving, (Englewood Cliffs, Prentice-Hall, 1972).
- 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, 226-230.
- 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, 4 (1977), 185-190.
- Michael J. Ryan, "An Examination of an Alternative Form of the Behavioral Intention Model's Normative Component," in H. Keith Hunt, (Ed.), Advances in Consumer Research, V (1978), 283-289.
- Herbert Simon and Allen Newell, "Human Problem Solving: The State of The Theory in 1970," American Psychologist, 26 (February, 1971), 145-159.
- 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 (1920), 25-29.
- Peter Wright, "The Use of Phased Noncompensatory Strategies In Decision Between Multi-Attribute Products," Research Paper 223, Graduate School of Business, Stanford University, August, 1974.
- , "Consumer Choice Strategies: Simplifying vs. Optimizing," Journal of Marketing Research, 12 (February, 1975), 60-67.

DECISION RULE UNCERTAINTY, EVOKED SET SIZE, AND TASK DIFFICULTY AS A FUNCTION OF  
NUMBER OF CHOICE CRITERIA AND INFORMATION VARIABILITY

Joseph J. Belonax, Jr., Western Michigan University

Abstract

If consumer brand choices are implemented in terms of decision rules, applying more choice criteria to information which contain more variability and a lower mean level should, (1) increase the uncertainty of constructing an effective decision rule, (2) increase the likelihood of achieving a smaller evoked set, and (3) decrease the difficulty in evaluating brands within a product class. Three 2 x 3 factorial experiments to test these hypotheses are reported. Results support the hypotheses.

Problem

Chaffee and McLeod (1973) present a generalized model that conceptualizes the consumer as an information processor and decision maker. The consumer is seen as making a buying decision according to some decision rule which combines information from an attribute-by-brand matrix. Within the context of this broad model, research on the decision rule aspect of information processing has taken many forms (summarized in Bettman, 1977).

Most often the studies involve fitting one or more decision rule models to brand evaluations made by persons who have been exposed to an attribute-by-brand matrix of information. The somewhat mixed results of studies attempting to fit various models may reflect individual and situational differences in those studies.

The consideration of which internal factors affect the choice of decision rule is well researched in the marketing literature. Park (1976) has considered prior familiarity. Wright (1975) considered the effect of such factors as the desire to simplify, and the desire to optimize choice strategies. Popielarz (1976) suggests that breadth of categorization as one dimension of cognitive style may also determine the relevant processing rule.

Previous investigations on environmental factors have focused on what effects the actual or perceived situation have on the choice of decision rule (Hansen, 1976). Studies on the actual situation have focused on such factors as, information load (Jacoby, Speller, and Kohn, 1974), task complexity (Reilly, and Holman, 1978), time pressure and distractions (Wright, 1974), structure and format of product information (Van Raaij, 1977). Effects of the perceived situation have centered on such variables as, perceived risk (Bettman, 1973), and perceived alternatives (Howard and Sheth, 1969).

Although individual differences may in part, account for mixed results of studies attempting to fit various models to brand evaluations, the distinction between the actual and perceived situation may be the more important factor to consider.

First, the distinction suggests that 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 from his awareness set (Narayana, Markin, 1975). The implied two process model, with information first used to form a set of perceived alternatives (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 the acceptable alternatives were considered.

Second, the distinction suggests that any attempt on the part of the consumer to construct a decision rule to evaluate brands may depend more on how the situation is perceived rather than on the actual stimuli present in the environment. As Bettman (1973) has pointed out in his theoretical model of inherent risk, the likelihood of constructing an effective decision rule depends upon the perceived distribution of quality over the brands in a product class.

Third, the distinction between perceived and actual situations is related to the distinction between anticipated and actual situations (Hansen, 1976). Since we believe that peoples' perceptions govern their behavior it is important to have measures for the before (perceived uncertainty) and the after (difficulty of the evaluation task).

In summary, individual factors are likely to influence consumer perceptions of the environment. In turn, these perceptions may influence the construction of a decision rule to evaluate brands, determine which of the brands will be perceived as acceptable alternatives, and ultimately affect the difficulty encounter in the evaluation task. This study will examine what effect the distribution of quality over brands in a product class and the number of choice criteria used to evaluate the brands have on the consumers' perceived uncertainty of rule construction, size of the evoked set, and degree of task difficulty.

Hypotheses

Although all consumers may be exposed to the same set of alternative brands in an experimental setting or in the market place, the distribution of quality over brands in a product class is likely to influence the consumers' perceived distribution of quality. According to the decision rule component in Bettman's (1973) inherent risk model, the greater the perceived variation in quality and the lower the mean level of quality distribution the less the likelihood of constructing an effective decision rule. Initially, consumers should feel less certain they can develop a rule to evaluate brands when the brands differ markedly in quality than when they differ very little. Further, Wright (1975) suggests that an anticipated increase in information load should increase the consumer's need to simplify his choice strategy. Initially, consumers applying more choice criteria to their awareness set should feel a greater need to engage in a simplified choice strategy. In turn, they should feel more certain that such a decision rule can be constructed.

Since the purpose of evoked set construction is to simplify the ultimate choice process, it seems likely that a person would eliminate brands which failed to meet some minimum level on one or more evaluative criteria. It would follow that, as the variability in quality increased and the general level of their desirability decreased, the likelihood of finding a reason for eliminating brands would increase. Further, as the consumer applied a larger number of evaluative criteria to the awareness set,

the likelihood of finding a reason for eliminating brands should be greater.

A greater variability of attribute values facilitates the actual differentiation of the choice alternative on that attribute. We may expect an increase in quality variability and decrease in the general level of desirability, to increase the likelihood of differentiating the brands on their attributes. This being the case, the difficulty in evaluating the brands should decrease. Further, consumers applying more evaluative criteria are likely to be engaged in a greater number of comparisons. In turn, their evaluation procedure should be more difficult. Therefore, specifically hypothesized in this study are the following:

- Hypothesis 1: The uncertainty of constructing a decision rule is inversely related to the number of choice criteria used in the evaluation task.
- Hypothesis 2: The uncertainty of constructing a decision rule is directly related to the variability and inversely related to the mean level of the attribute ratings presented in the evaluation task.
- Hypothesis 3: The size of the evoked set is inversely related to the number of choice criteria used in the evaluation task.
- Hypothesis 4: 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.
- Hypothesis 5: The difficulty in evaluating the brands is directly related to the number of choice criteria used in the evaluation task.
- Hypothesis 6: The difficulty in evaluating the brands 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, three 2 x 3 factorial experiments were designed with "number of evaluative criteria" and "variability and mean level of presented attribute information" as independent variables and the reported uncertainty of constructing a decision rule, size of the evoked set, and task difficulty as the dependent variables. The preliminary design considerations are discussed in detail elsewhere (Belonax and Mittlestaedt, 1977). However, they will be presented briefly before the operationalization of the variables are discussed.

First, micro-wave ovens were selected because the product is sufficiently complex to allow the use of several evaluative criteria, and one in which participants would be somewhat interested.

Second, to insure that attributes of importance were selected the "direct dual questioning technique" described by Alpert (1971) was used. The six attributes selected included browning uniformity, cooking precision, cooking uniformity, cooking versatility, oven capacity and warranty protection.

Third, to manipulate the variable of the attribute ratings "Standard squares" were used. Since there were six attributes, information was presented on six "brands" to each participant.

Fourth, control for product familiarity was achieved by selecting 300 students in an introductory nutrition course which had not discussed micro-wave ovens in class. Participants were randomly assigned to treatment levels. 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 held constant within treatment levels, and systematically manipulated between treatment levels. The mean level and variability are from the "low variability" condition ( $\bar{X} = 8.0$ ;  $s^2 = 5.9$ ), through the "intermediate variability" condition ( $\bar{X} = 7.0$ ;  $s^2 = 8.4$ ), to the "high variability" condition ( $\bar{X} = 6.0$ ;  $s^2 = 14.0$ ).

The other independent variable was a stratification according to the number of choice criteria used. After examining the brand ratings the participants were asked to indicate: (a) how certain they were of constructing a decision rule on a twelve-point scale, from "extremely certain" 1 to "extremely uncertain" 12. (The answer to this question is the first dependent variable). After making an overall evaluation for each brand, participants were asked to indicate: (b) the importance of each attribute to the evaluation decision, (c) the minimum rating on each attribute which an oven would have to receive to be evaluated as "acceptable", (d) the number and designation of brands considered to be acceptable for purchase, (The answer to this question is the second dependent variable), (e) how difficult it was to evaluate the brands on a twelve-point scale, from "extremely easy" 1 to "extremely difficult" 12. (The answer to this question is the third dependent variable). Answers from questions (b) and (c) were combined and a mean calculated for each attribute. Any participant's combined answer on each attribute was designated "employed choice criteria" 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.

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 ANALYSIS (Finn, 1972). Although there are some differences in cell size, reordering the variables for unequal cells did not affect significance.

#### Results

Table 1 shows the mean level of uncertainty for the various conditions.

An examination of the means shows that, as predicted by Hypothesis 1, those employing "more" choice criteria were more certain of constructing a decision rule. Hypothesis 2 predicted that uncertainty would increase as the variability and mean level of attribute increased

TABLE 1  
Mean Level Of Uncertainty For All Conditions

Choice Criteria	Attribute Rating Variability			Total
	Low	Intermediate	High	
Many	4.91 (58)*	5.18 (38)	6.35 (37)	5.39 (133)
Few	5.15 (44)	5.91 (60)	6.90 (63)	6.09 (167)
All	5.02 (102)	5.63 (98)	6.70 (100)	6.24 (300)

\* ("n" in parentheses)

and decreased, respectively. The means are in the predicted direction for both the "many" and "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 Hypotheses 1 and 2.

TABLE 2  
Univariate Analysis Of Variance

Source of Variation	d. f.	Mean Square	F
Choice Criteria	1	36.158	6.812 (p < .01)
Rating Variability	2	63.883	12.034 (p < .01)
Interaction	2	1.473	.278 (NS)
Error	294	5.308	

Table 3 shows the mean evoked set size for the various conditions of the study.<sup>1</sup>

TABLE 3  
Mean Evoked Set Size For All Conditions

Choice Criteria	Attribute Rating Variability			
	Low	Intermediate	High	Total
Many	2.55	2.31	1.83	2.29
Few	3.09	2.78	2.14	2.62
All	2.78	2.60	2.03	2.46

An examination of the means shows that as predicted by Hypothesis 3, those employing more choice criteria constructed smaller evoked sets. Hypothesis 4 predicted that the evoked set would decrease between treatment levels. The means are in the predicted direction. Table 4 presents the results of the two-way univariate analysis of variance. Both main effects are significant while the interaction is not. The results confirm Hypothesis 3 and 4.

TABLE 4  
Univariate Analysis Of Variance

Source of Variation	d. f.	Mean Square	F
Choice Criteria	1	8.410	8.987 (p < .01)
Rating Variability	2	18.286	19.541 (p < .01)
Interaction	2	.344	.368 (NS)
Error	294	.936	

1 Results of this analysis are presented elsewhere (Belonax and Mittlestaedt, 1977).

Table 5 shows the mean task difficulty for the various conditions.

The means show that as predicted by Hypothesis 5, those employing more choice criteria reported more task difficulty. Hypothesis 6 predicted that task difficulty would decrease between treatment levels. The means are in the

TABLE 5  
Mean Level Of Task Difficulty

Choice Criteria	Attribute Rating Variability			
	Low	Intermediate	High	Total
Many	6.67	6.37	5.89	6.37
Few	6.31	5.50	5.46	5.70
All	6.52	5.84	5.62	6.04

predicted direction. Table 6 presents the results of the two-way univariate analysis of variance. Both main effects are significant while the interaction is not. Hypotheses 5 and 6 are confirmed.

TABLE 6  
Univariate Analysis Of Variance

Source of Variation	d. f.	Mean Square	F
Choice Criteria	1	33.019	7.120 (p < .01)
Rating Variability	2	16.476	3.553 (p < .05)
Interaction	2	1.822	.393 (NS)
Error	294	4.637	

#### Discussion

Although no attempt was made to determine the separate effects of attribute variability and mean level of attribute quality, the results confirm the view that taken together they engender greater uncertainty on the part of the participants. The uncertainty is less when it is anticipated that many criteria will be applied to information. This suggests at least two observations.

First, that people appear to exhibit greater uncertainty dependent on mean level and variability of attribute quality implies that greater attention should be paid this phenomenon in information processing studies.

Second, although no model was fitted to preference orderings, the anticipated increase in information load brought about by the greater number of attributes to be considered per brand appears to increase the relative urgency of simplifying their choice strategy. Thus, it is likely that consumers anticipate using simplification heuristics in situations of perceived information overload and felt more certain of executing the strategy.

That the size of the evoked set seems to depend on the number of choice criteria used and the nature of available information is consistent with the expectation that participants would engage in a simplification strategy to construct an evoked set. 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 valid and/or less reliable

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 salience of the attributes presented and the variability and mean levels of the information to which participants are exposed.

The results confirm the view that, task difficulty appears to depend on the conditions of this study. This is consistent with the expectation that participants find it easier to differentiate brands that differ markedly in quality than brands that differ little. Further, it seems that the use of more criteria increases the actual difficulty of the evaluation task.

#### References

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

Joseph J. Belonax, Jr. and Robert A. Mittlestaedt, "Evoked Set Size as a Function of Number of Choice Criteria and Information Variability," in H. Keith Hunt, ed., Advances in Consumer Research, Vol. 5, Proceedings of the 8th Annual Conference of the Association for Consumer Research, 1978, 48-51.

James Bettman, "Perceived Risk and Its Components: A Model and Empirical Test," Journal of Marketing Research, 10 (May 1973), 184-90.

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.

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.

Fleming Hansen, "Psychological Theories of Consumer Choice," Journal of Consumer Research, 3 (December 1976), 117-142.

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

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 1974), 33-42.

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

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.

Donald Popielarz, "An Exploration of Perceived Risk and Willingness to Try New Products," Journal of Marketing Research, 4 (November 1967), 349-54.

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

Michael Reily 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.

Peter Wright, "The Harassed Decision Maker: Time Pressure, Distractions, and the Use of Evidence," Journal of Applied Psychology, 59 (October 1974), 555-61.

Pete R. Wright, "Consumer Choice Strategies: Simplifying Vs. Optimizing," Journal of Marketing Research, 12 (February 1975), 60-67.

Fred Van Raaij, "Consumer Information Processing for Different Information Structures and Formats," 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, 176-184.

James H. Myers, University of Southern California

While all three papers in this session expand upon our limited knowledge of evoked sets, they do so in very different ways. May theorizes a linkage between learning theory and information processing theory in evoked set formation based on parts of McGuire's information processing paradigm; Parkinson and Reilly utilize one aspect of consumer information processing theory as a basis for studying evoked set formation in an empirical setting. Both of these papers purport to focus primary attention upon the dynamics of evoked set formation. Belonax studies yet another set of factors that affect the size (but not composition) of the evoked set.

While all papers add to our knowledge in this area, there is some question that any of the papers, or any other work to date for that matter, provides any real insight as to how evoked sets are formed over time within an individual consumer. Parkinson and Reilly deal briefly with this issue when they state, "In an actual temporal sense what probably happens is that the consumer considers one brand at a time and makes a binary decision as to whether or not the brand is worthy of future consideration." This sounds reasonable, but is it really what happens? If it is, would any of the following three processing models be appropriate to reflect this one-at-a-time evaluation: weighted/unweighted linear compensatory, lexicographic? It would seem that these models in particular would be more appropriate for an evoked set formation process which considers all brands in an awareness set at the same point in time. However, it is entirely possible that evoked sets are formed over time by considering each new entry into the awareness set whenever it comes along. It would seem that a conjunctive or a disjunctive processing model would be much more appropriate for this type of evaluation, yet these models did not do nearly as well in predicting "actual evoked sets" as did the other three in the Parkinson and Reilly study. Does this mean that study results suggest that evoked sets are actually formed by comparing and evaluating all brands in the awareness set at a single point in time? But this does not seem to agree with the earlier statement by Parkinson and Reilly about brands being considered "one at a time," unless the entire existing evoked set is reevaluated each time a new entry comes along.

The point of this matter is not whether or not these authors are correct and consistent. The question is whether or not we really know much about the temporal aspects of evoked set formation: not only what kinds of information processing takes place but in what sequence this takes place and over what periods of time. And, of course, how this varies by such factors as psychological and social considerations, product complexity, the learning that might take place, and so on.

#### Need For a Better Methodology

Thus, while two of the papers emphasize the "dynamics" of evoked set formation, they do so within a static context. May considers the dynamics of the formation primarily by relating this to two concepts: stages in the adoption process, and stages in the product life cycle. While these factors are indeed dynamic in nature, they seem more appropriate for aggregates than for individuals. There is certainly nothing wrong with examining individuals at different points within each cycle for clues as to what kinds of processing are going on, and this is what May has done. Within the static context,

both papers provide some new insights in an area not very well understood.

However, my own view is that we need: (1) a better theory to describe what is going on over time as a given consumer forms his/her own evoked set, (2) research designs that are longitudinal rather than cross-sectional, to trace the formation process within individuals who are in different stages of the process.

For example, it would not be difficult to trace the formation process over time within college students or mail panel participants. The research design might call for the selection of some product category that is new (entirely or relatively) to each respondent and also of some degree of interest. This would be decided by each person on an individual basis. An inventory of the awareness set, the evoked set, and even inert and inept sets (Narayana and Markin, 1975) could be taken at 1 or 2 month intervals over a designated time period. The respondent would be encouraged to learn as much as possible about the product category, using whatever sources of information and/or trial he/she chooses in the normal manner.

At each inventory period, information is gathered about as many aspects of the evoked set formation as possible: present awareness set; evoked, inept and inert sets; sources of first and subsequent information about each new brand entering the awareness set; trial; attribute ratings for familiar brands; preference rankings, etc. The resulting data could be analyzed to track the movement over time of brands into or out of awareness, evoked, inept and inert sets. This movement could be correlated with such factors as current numbers of brands in each of the sets, attribute ratings, information sources, and many others.

This type of longitudinal data would make it possible to construct temporal information processing networks showing the real dynamics of establishing each of the types of sets, the stability of these sets over time, and the types of factors that appear to have the greatest influence on formation, size and composition of the sets. The data could also be tested against the May paradigm combining the information processing and learning theory approaches to evoked set formation. The various brand-attribute ratings could also be processed by the five models used by Parkinson and Reilly to shed light on why a particular brand went into or out of any of the sets.

Another possible research design would follow the suggestions of Parkinson and Reilly that consumers be presented with "the introduction of a totally new brand for which there is no prior evaluation." This brand would presumably be in the form of a detailed concept statement. Several of these concepts could be introduced to the same individuals at successive time periods, and the resulting information processing could be monitored in the manner described above. For this design perhaps all respondents would be asked to consider the same familiar product category, such as toothpaste or deodorant.

#### Methodological Considerations

If the above longitudinal information is to be analyzed using either the May or the Parkinson and Reilly design, there are a few methodological improvements that should be considered. It would seem that May could provide bet-

ter operational definitions for some of the cells in his model; e.g., types of information source users, search and retrieval constructs. It would also be helpful if he expanded upon his taxonomy of evoked sets to include considerations other than trial. Other investigators have dealt with size and composition of evoked sets, and it would seem that these factors could and should be studied in longitudinal designs.

Parkinson and Reilly have pointed out their own potential problems with halo and especially with attempting to measure both conjunctive and disjunctive cutoffs. Belonax has used their same approach to establishing cutoffs and would seem to have this problem also. This problem in particular must be solved if we are to have any meaningful test of the congruence between these models (conjunctive and disjunctive) and evoked set formation. My own suggestion is to ask respondents to rate all brands in the awareness set on each attribute and then to indicate brands that are either unacceptable or extremely acceptable on that attribute. Tying cutoffs to actual brands for each attribute should help in clarifying the different levels of acceptability and in identifying brands that either are or are not in the consideration set.

If conjunctive and disjunctive cutoffs were specified more accurately we could then investigate the possibility of a phasing model, which might postulate the establishment of an evoked set by one of these cutoff models followed by evaluation of each brand within the set using a compensatory model of some sort.

The general notion of phasing seems more appropriate to the dynamics of evoked set formation over time.

#### Buying Consideration

One of the problems I see in most studies of evoked sets is the matter of operational definitions of the various types of sets. In particular, the evoked set is defined by most investigators (and by Howard & Sheth) as brands the consumer would "consider buying." But Belonax asked what brands the participant would consider to be "acceptable for purchase." Do these terms mean the same thing? And do they mean the same thing for soft drinks as for automobiles? Are there brands that a consumer would consider acceptable but would not seriously consider buying? I believe this is entirely possible.

In the case of automobiles or other expensive items, it might be that several brands would be within some sort of "acceptable" or "possible" range for a consumer but that only one or two brands would be "seriously considered" due to a risk aversive purchase strategy which requires prior experience with the brand. In the case of soft drinks there is little risk for any purchase, and consumers often consider many different brands for reasons of variety or because each brand may be most appropriate for a certain time of day or drinking situation.

This is more than simply quibbling over semantics. At present the evoked set is considered monolithic. We need construct analysis studies to tell us if there is more than a single layer within this set. It might be that "consider buying" is at least a continuum or perhaps is multidimensional. This would add greatly to our understanding of the evoked set construct and it would certainly provide guidance for designing future studies on the formation of these sets.

#### References

- Joseph J. Belonax, "Decision Rule Uncertainty, Evoked Set Size, and Task Difficulty as a Function of Number of Choice Criteria and Information Variability," Paper presented at Association of Consumer Research Annual Meeting, October, 1978.
- Chen L. Narayana and Ron J. Markin, "Consumer Behavior and Product Performance: An Alternative Conceptualization," *Journal of Marketing*, 39 (October, 1975), 1-6.
- Thomas L. Parkinson and Michael Reilly, "An Information Processing Approach to Evoked Set Formation," Paper presented at Association for Consumer Research Annual Meeting, October, 1978.
- Frederick E. May, "Evoked Set Formation and Composition: The Learning and Information Processing Hypotheses," Paper presented at Association for Consumer Research Annual Meeting, October, 1978.

TABLE 1  
Types of Attitude Scales  
in Consumer Research Studies

Scale and Validation	No. of studies using	% <sup>a</sup>
<u>Scale Type</u>		
Multi-attribute	60	47
True Likert	1	1
Likert type	31	24
True semantic differential	2	2
Semantic differential type	23	18
Rank order preference <sup>b</sup>	21	17
"Agreement scale"	3	2
Happy/smiling faces	2	2
"Liking scale"	2	2
"Bipolar scale(s)"	2	2
"Favorableness scale"	2	2
15-item equal appearing interval	1	1
Stapel	1	1
Decision net	1	1
"7-point scale"	1	1
"Appeal scale"	1	1
Could not determine	19	15
<u>Validation Attempted?</u>		
Yes	43	34
No	84	66
Studies in which "true" validation used	3	2

<sup>a</sup>Some studies included two or more types of scales, usually comparing them to each other or some criterion or validating measure. Thus, percentages add to more than 100.

<sup>b</sup>Includes paired comparison preference measures.

The large number of articles, 127, in which attitude measures were utilized precludes study-by-study review. Instead, some examples will be discussed. These examples are not intended to embarrass the researchers, indeed this writer is as guilty as any.

The difficulty of developing a Guttman scale was demonstrated by Pessemier, Bemmaor, and Hanssens (1977). They were forced to reject their hypothesis that their items, concerning willingness to donate body parts, could be scaled in one dimension. Admirably, they not only admitted this problem, but also included ten "attitude scales" as independent predictors of willingness to donate. Each scale, consisting of four items each, was Likert-like, but the attitude constructs seemed unlikely to be tied, directly or indirectly, to a willingness to donate body parts. Instead, such scale items as "Every woman should be free to have an abortion" and "I try to impress people with my leadership ability" seem more related to psychographics/life style research. There is no quarrel with the researchers' relational findings, but neither is there a validated attitude measure in their article.

Hawkins, Albaum, and Best (1974) provided a test of the Stapel scale using the semantic differential as a validating construct. Unfortunately, their semantic differential scales had six rather than seven scale positions and included such nonstandard items as "helpful employees-unhelpful employees" and "limited selection-wide selection." Also, no internal validation was provided to develop a semantic differential attitude measure. Hawkins, *et al.*, did not claim to be testing an alternate attitude measure, however, but an alternate to

the semantic differential because they felt the Stapel scale was easier to administer in telephone interviews. Appropriate replication of this study testing for Stapel-semantic differential attitude scale equivalence would be useful. Such studies should include a semantic differential telephone administration treatment which was not included among the original treatments.

In an early consumer research test of the Fishbein extended model, Bonfield (1974) utilized a multi-attribute model described by Sheth (Howard and Sheth, 1969; Sheth, 1969). Since the Sheth model is nonstandard, it should have been validated. No report of a validation attempt was shown. More recently, Holbrook (1977) has used a "nine-position general liking scale (p. 168)" with "do not like at all" and "like very much" as endpoints as a criterion variable for a multi-attribute attitude model. While the criterion scale has face validity as a measure of affect, it suffers, nonetheless, from a lack of validation against a standard scale. Moreover, as a one-item criterion measure, it is impossible to test for internal reliability. More seriously, perhaps, the criterion scale may suffer from ceiling effect on the positive side at least. That is, among well-liked singers, all would likely have top scores among individuals, even though their actual attitudes would not be equally positive. Bonfield and Holbrook are only two examples of researchers who have used sophisticated theoretical and mathematical concepts, such as Holbrook's optimally scaled model, yet failed in the critical validation aspects of the models they used.

If, as is contended with the examples discussed, most consumer research attitude measures have not been shown to measure attitude, how should the researcher go about checking the validity of nonstandard attitude measures?<sup>3</sup> The development of a standard scale measuring attitude toward the same object, concept, or behavior as the non-standard scale is one viable strategy. The standard measure acts as an independent measure of concurrent and/or convergent validity (Zaltman, *et al.*, 1975, pp. 39-47). Because Thurstone, Guttman, and Likert scales are time consuming, and sometimes difficult, to construct, semantic differential attitude scales are recommended as a validating construct. Since most multi-attribute models have been operationalized utilizing semantic differential type scales, convergent validity is sacrificed when semantic differential methodological protocols are utilized. Following an appropriate research protocol, semantic differential scales are easily selected (i.e., no development is necessary), easily administered, and, with computers and factor analysis programs, easily analyzed.

It should be recognized, however, true semantic differential scales utilize standard instructions which can be paraphrased, have seven places for each scale-item, and only bipolar scale item names selected from a specific set can be used. This information can easily be found in Osgood, Suci, and Tannenbaum (1957). Some 10 to 20 scales should be selected, most associated, a priori, with the evaluative dimension, but with some associated with the activity and potency dimensions of meaning. Once data is collected, factor analysis--actually principal components analysis with varimax rotation--will reveal those scales with the highest loadings on the evaluative dimension, the only semantic dimension representing attitude. If high enough loadings are found for three to five evaluative scales, scores on these scales can be summed or averaged for each individual as an attitude criterion variable. High factor loadings can be taken as indicants of scale reliability rendering additional tests for reliability in the criterion measure superfluous.

<sup>3</sup>A number of sources can be found which discuss validity more thoroughly than is possible here. One such source, Zaltman, Pinson, and Angelmar (1975, pp. 39-47) is directly tied to consumer research.



The reasons for imperfect attitude-behavior relationships are, for the most part, fairly obvious. First, attitude is only one of a number of factors influencing behavior. Other factors, such as personality or social influence, may act so as to block behavior that attitude would favor. Much social psychological and consumer research has been expended on the Fishbein extended model (see Fishbein and Ajzen, 1975; Ryan and Bonfield, 1975) which is described as a method for studying attitude in relation to other factors as they relate to behavior.

Secondly, attitudes may be weakly held and thus easily changed. In those cases, even if attitude is a good predictor of behavior, only when the attitude and behavior measurements are nearly simultaneous could we expect a strong attitude-behavior relationship to be found.

Third, attitude may, in some cases, not even exist. Since most respondents will provide information on attitude scales unless they have absolutely no knowledge of its subject, it should not be surprising when "attitudes" among these respondents have little or no relationship to their behavior.

Fourth, behavior with respect to an object, concept, or act can be viewed as learned after the attitude is learned (Doob, 1947, as described by Fishbein, 1967b, p. 478). That is, attitude is treated by these theorists as a learned mediating response (Doob, 1947; Rhine, 1958; Fishbein, 1967a,b; Olson and Mitchell, 1975). Once attitude is learned, the individual must also learn what response to make to that attitude.

...that is, there is no innate relationship between the attitude and behavior; one still has to learn a behavioral response. Two people may learn to hold the same attitude toward a given stimulus; clearly, however, they may also learn to make different responses given the same learned attitude.

For example, two students may learn to feel equally favorable toward a given instructor. Furthermore, this feeling may initially elicit the same overt response in both students (e.g., calling the instructor by his first name). From Doob's point of view, the probability that this behavior will persist is a function of the reinforcement the students get for making this response. For example, with respect to one student, the instructor might say, "Well, I'm glad you finally decided to drop that 'professor' nonsense," while he might tell the other student, "I'd prefer it if you wouldn't call me by my first name." If this were the case, Doob's theory would predict that the behavior would continue for the first student, but not for the second (Fishbein, 1967b, p. 478).

Fishbein recognized that, in the given example, the non-reinforced behavior would probably lead to attitude change as well as differential behavior.

Fifth, since attitude is a predisposition to act rather than an act itself, it is possible to view attitude as a wish or impulse. In that sense, attitude may be closely related to the Freudian concept of the id, which, according to Hall and Lindzey (1968), "is able to do only one thing, namely, to wish (p. 249)." If so, and since the wishes of the id are monitored and frequently modified by the ego and superego, not only would we expect much less than perfect attitude-behavior relationships, but we should also expect positive attitudes toward many "socially unacceptable" objects, concepts, and behaviors. This notion suggests indirect measures of attitude, such as the semantic differential, will be most accurate as attitude measures because they are less subject to ego and superego moderating effects which would show up as more socially acceptable attitudes. Alternately, it may be possible individuals have three potentially distinct attitudes with respect to each object, concept, or act;

one each associated with the id, ego, and superego.

For other reasons attitude may not predict behavior, see Cohen (1974, pp. 341-343) and Cohen and Ahtola (1971, p. 346). If behavior is not a validator of attitude measures, a fact tacitly recognized by most consumer researchers, how then are we validating our attitude measures?

#### Validation of Nonstandard Attitude Measures

Consumer researchers have utilized a number of attitude measures, many of which are nonstandard in the sense of not being Thurstone, Guttman, Likert, or semantic differential attitude scales. In particular, consumer researchers appear enamored with multi-attribute attitude models.

A limited review of the literature reveals the reason for the title of this paper. The review includes three sources: (1) all issues of the Journal of Marketing Research from February 1974 to February 1978 which loosely defines a "post-Wilkie-Pessemier (1973) era" for attitude, particularly multi-attribute attitude measurement; (2) all issues of the Journal of Consumer Research which was first published in 1974; and (3) all Association for Consumer Research Conference Proceedings. An item had to be available April 1, 1978 to be included in the review.<sup>2</sup> Some double counting was included since multiple publications of the same data set, except in review articles, were treated as separate studies. The results of the data review are summarized in Table 1.

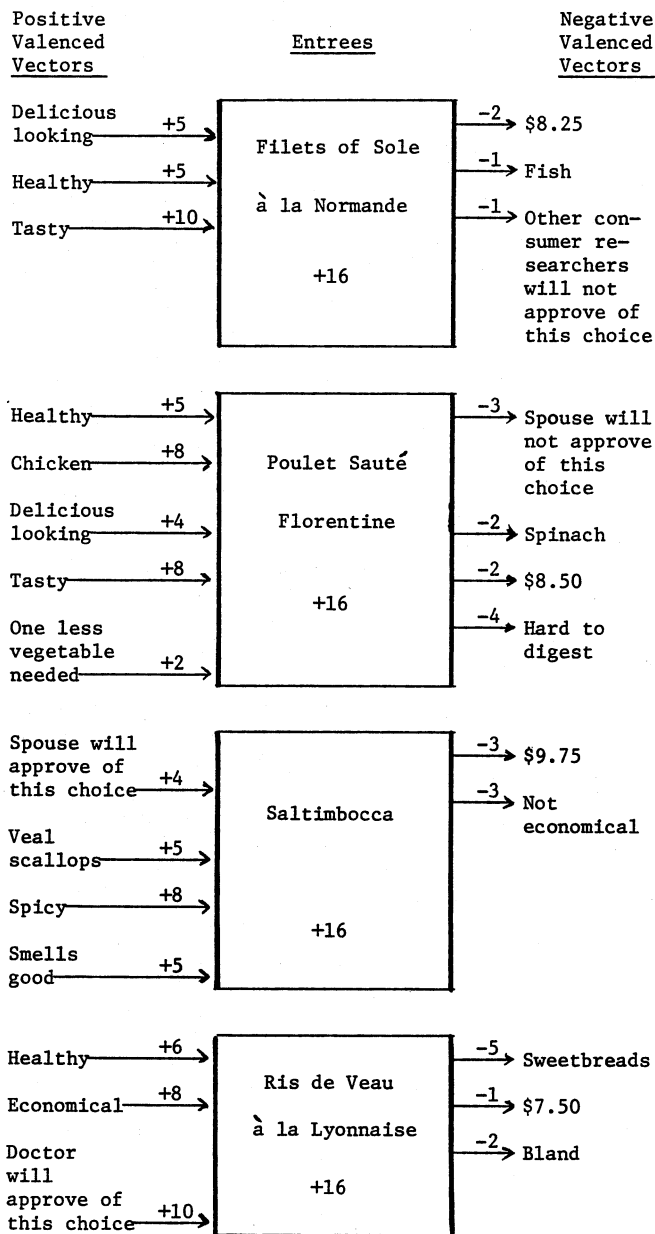
As can be seen, nearly half the studies utilized scales which were operationalizations of some form of multi-attribute model. There were three measures for which enough information was presented for classification as a standard attitude measure, although in a number of cases the technique could be considered an approximation of a standard scale. Few researchers stated that any attempt was made to validate any attitude scale, but in many cases, terms such as "criterion" and "independent" measure were used in such a way as to suggest high correspondence between the two measures would be acceptable evidence of validity. Of these, only Lutz (1972, 1975) used a true standard measure, semantic differential scales, in such a way as to test the validity of a multi-attribute measure. In the more recent study, Lutz (1975) was treating the multi-attribute model as an indicant of the underlying cognitive structure of attitude rather than a measure of attitude. While Lutz' study is classified as a multi-attribute study with a true validation attempt in Table 1, it is also clear the study could have been classified as one in which a true semantic differential attitude measure was used. At any rate, it is clear that Lutz, for one, knows how to use the semantic differential to measure attitude since in the earlier study, Lutz (1972) had tested a multi-attribute model, called an attitude model, against an appropriately developed semantic differential measure of attitude.

Some researchers (e.g., Arndt and Crane, 1975) specifically addressed themselves to the limitations of their measures while others (e.g., Cox, 1975) did not specify that their operationalizations were attempts to measure attitude. Finally, a number of researchers related their attitude measures to preference ranks. While perhaps related to Thurstone's law of comparative judgment, preference ranks have not been substantiated as an attitude validating construct.

<sup>2</sup>The purpose of this review was to document a trend. Since it is inappropriate to enter into data search with biased expectations, it is hoped some degree of scientific objectivity was maintained throughout the data search.

FIGURE 1

A HYPOTHETICAL RESTAURANT ENTREE EXAMPLE OF APPROACH-APPROACH CONFLICT



positive net valence greater than all others, that entree would be easily chosen, although some conflict would still be felt since the positive valence associated with unchosen alternatives would be lost and, at the same time, the negative valence associated with the chosen alternative would be obtained along with the positive valence.

This example has two major implications for understanding why multi-attribute models are not good predictors of behavior. First, in Figure 1, each choice has a unique set of associated vectors and no one vector name is associated with all choices—clearly a case of heterogeneity of salient attributes. Entrees on a menu, it can be argued, are different from brands in a product class. However, it is contended the difference is one of number of heterogeneous vectors rather than whether heterogeneity exists. For example, the vector "Superstar in Rent-A-Car" cannot, or more accurately, should

not, be associated with Avis, although "Not Superstar in Rent-A-Car" can. "Marlboro Country," or even "Manly Cigarettes," should be a frequently found vector for Marlboro cigarettes unlikely to be found for Winston cigarettes. The fact is much marketing effort is expended to create these unique vectors which are called differential advantages, yet nearly every study which has utilized multi-attribute models across brands in a product class has assumed each brand has exactly the same set of attributes/vectors as every other brand in the product class.

The second implication shown by the example depicted in Figure 1 is that not all vectors are attitudinal or evaluative in content. Some vectors represent values, e.g., healthy, economical; some represent social influence, e.g., spouse will approve, doctor expects me to order this; some represent definitional aspects, e.g., chicken, veal scallops, \$8.25; and some are clearly evaluative, e.g., delicious looking, spicy, tasty. The second implication leads to a dilemma: (1) Should researchers concentrate on multi-attribute models in which only attitudinal vectors are considered in attempts to measure attitude as one of many influences on behavior? (2) Or, should researchers attempt to develop more global multi-attribute models which include all salient vectors in order to predict and understand behavior, but not attitude? The bulk of consumer research thus far appears to have been attempts to answer the first question.

It is hypothesized research aimed at answering the second question will lead to two results. First, more global multi-attribute models will lead to better prediction of attitude. This expectation follows from findings that attitude is not independent of other influences on behavior. For example, in their review of research undertaken following the Fishbein extended model, as well as in their further joint research, Ryan and Bonfield (1975, 1977) have shown evidence that while attitude and social influence provide independent explanations of variance in intentions, a joint, or interaction, effect is also present.

Secondly, more global multi-attribute models will lead to better predictions of intentions and possibly behavior than, for example, Fishbein's extended model. Bither and Wright (1977) have utilized a global multi-attribute model to explore preferences for information sources. An interesting re-analysis of their data might test the homogeneity assumptions. A model and procedure suggested by Bernardo and Blin (1977) also appears to have potential value as a global model, but seems to rest on a homogeneity of attributes assumption. Stanton and Lowenhar (1974) have defined a need-press model—in which press is treated as an alternative's ability to satisfy needs—which was operationalized as a multi-attribute model. While the Stanton and Lowenhar model was not a predictor of rank order preference, relative preference, or attitude as measured using a Thurstone scale, it was not tested as part of a global multi-attribute model for the predictions of intentions or behavior. It is assumed researchers would continue to attempt to find optimum combination rules for data collected following a more global model.

Such a model may extract a cost, however, in terms of understanding. That is, a global multi-attribute model may obscure relationships among attitudes, social influences, needs, and other influences on behavior. At the same time, more global, multi-attribute models with appropriate identification of vector-attributes can have strategy directing, strong relationships with behavior which are not devoid of potential for understanding behavior. In that sense, such a model—which is not an attitude model—would be in line with Kassarian's (1978) call for simpler models.

Since attitude structure may be different across objects, concepts, and acts, as well as across respondent groups, separate factor analyses should be performed for each regardless of the original set of scales selected.

While the accurate adherence to a standard measure protocol can be considered a validation procedure in and of itself, even standard scales should be checked for validity. Lundstrom and Lamont (1976), for example, checked both internal and external validity of their Likert-type scale designed to measure consumer discontent. Validity testing included sampling from two groups expected to hold favorable, i.e., contented, and unfavorable, i.e., discontented, attitudes. The former group consisted of members of Rotary and Kiwanis, business and professional organizations, while the latter group consisted of members of a consumer activist organization and people who had lodged complaints with a Better Business Bureau. Internal validity testing consisted of correlating item scores with total scores while external validity was tested by comparing attitude scores of the two groups. Testing external validity using known groups is also a viable strategy for non-standard attitude measures.

Baker and Churchill (1977) utilized semantic differential-like methodology in an apparently successful attempt to measure affective, cognitive (belief), and conative (intention) dimensions of reactions to advertisements. Rather than factor analysis, however, these researchers utilized item-to-total correlations, thus creating a need for a validating measure which was not included. Other researchers (e.g., Barnes, 1978) have also attempted to tap these three dimensions.

It would be impossible to exhaust validation possibilities in this paper. Hopefully, the argument that validation is both needed and possible can be accepted by consumer researchers interested in measuring attitudes.

#### Why Multi-Attribute Models Don't Work

Actually, appropriate multi-attribute models may "work." First, one or more forms of the model may be a valid measure of attitude, but appropriate evidence supporting such a model has not been presented. Even if they cannot measure attitude, appropriate multi-attribute models may be valuable as indicants of viable strategies for changing attitudes and/or behavior through their presentation of a picture of the cognitive structure underlying attitude. At the present time, we have scant knowledge of which operational form of such a model, or which statistical manipulation of the data collected, if any, is a valid measure of attitude. There are, however, reasons to expect weak relationships between multi-attribute models and criterion, standard attitude measures.

Specifically, two homogeneity assumptions are usually, tacitly, made relative to the operationalization of multi-attribute models. These assumptions are (1) all individuals will have the same set of attributes underlying their attitudes and (2) all objects within a class of objects (e.g., brands within a product class) will have the same attributes. Conversely, the problems of heterogeneity of salient attributes across groups of people and heterogeneity of attributes across objects, concepts, or behaviors within classes are simply not addressed. The problem of determining salient attributes is not addressed here. For discussions of that problem see Ryan and Bonfield (1975, pp. 121, 122) and Ryan and Etzel (1976).

Homogeneity of salient attributes within groups should be a reasonable assumption, otherwise it would probably be useless to deal in multi-attribute models. Nonetheless, since, as Wilkie and Weinreich (1972) have shown,

perfect homogeneity of salient attributes across members within a group is unlikely, multi-attribute models will be relatively poor predictors of attitude as measured using a standard procedure. In addition, no case was found in which the homogeneity assumption within groups was tested. Any researcher desiring to test this assumption should be most careful in selecting variables which differentiate groups. Inappropriate differentiating variables will tend to magnify within group heterogeneity.

A more serious error is to treat all members of all groups as having the same set of salient attributes, i.e., a problem of heterogeneity of salient attributes across groups. Upper income group members are less likely, for example, to regard price as a salient attribute of sirloin strip steak than low income group members. Assuming product knowledge, it is likely both groups would rate the object the same on multi-attribute scales even though the price scale(s) would have no bearing on the upper income groups' attitudes. Only Ryan and Etzel (1976) appear to have tested for heterogeneity of salient attributes across groups and across brands. Cox (1975) assumed heterogeneity of salient attributes, but did not test this assumption.

The likelihood of heterogeneity of salient attributes across objects in a class may not be so obvious. For demonstration purposes, it is useful to turn to Lewinian conflict theory as discussed in almost any basic psychology text. Lewin's field theory is concerned with choice. In any choice situation, the individual is motivated to move from one state which is negatively valued, say "hungry," to some other state which is positively valued, "not hungry," for example. Other motives will be operative at the same time, for example, to move from "thirsty" to "not thirsty" and from "need to write manuscript" to "finished manuscript." In moving from a state of "hungry" to "not hungry," the individual may choose from a vast array of means, called pathways in field theory. Any one pathway may not be mutually exclusive compared to other pathways, just as several highway routes from Chicago to Dallas may require driving part of the way on the same stretch of highway as other routes.

Suppose the individual selected a pathway which included eating in a restaurant. The selection of which restaurant follows a decision to eat at a restaurant rather than a fast food franchise, a delicatessen, or at home, among others, of course, but for our example, it will be useful to have our individual at the point of selecting an entree at the chosen restaurant. Each item on the menu has associated with it a number of vectors, which in terms of the present example, are defined as field forces leading toward or away from a specific choice. In Figure 1, vectors are represented by arrows. Each vector has a property called valence which by sign, plus or minus, indicates the direction of the vector; and by degree of positiveness or negativeness, indicates the strength of the vector. Strength of each vector is indicated by the numerical value placed next to the signs in Figure 1. Obviously, a vector is analogous to an attribute. Figure 1 is an example of approach-approach conflict, that is, all the relevant choices have equal and positive net valence.<sup>4</sup> If any entree choice had a

<sup>4</sup>Other field theory conflict types are not treated here. In avoidance-avoidance conflict, all relevant choices have equal and negative net valences. With ambivalence, the choice is whether to perform a particular, general behavior rather than which of two or more behaviors to perform. Thus, an individual may be ambivalent as to whether to buy a color television set. Ambivalence occurs when the positive and negative valence associated with the vectors cancel each other out leaving a net valence of zero.

Consumer Research, 1, March, 49-59.

Olson, Jerry C. and Andrew A. Mitchell (1975), "The Process of Attitude Acquisition: The Value of a Developmental Approach to Consumer Attitude Research," in Advances in Consumer Research, Volume 2, ed. Mary Jane Schlinger, Association for Consumer Research, 249-64.

Osgood, Charles E., George J. Suci, and Percy H. Tannenbaum (1957), The Measurement of Meaning, Urbana: The University of Illinois Press.

Pessemier, Edgar A., Albert C. Bemmaor, and Dominique M. Hanssens (1977), "Willingness to Supply Human Body Parts: Some Empirical Results," Journal of Consumer Research, 4, December, 131-40.

Rhine, Ramon J. (1958), "A Concept-Formation Approach to Attitude Acquisition," Psychological Review, 65, 362-70.

Ryan, Michael J. and E. H. Bonfield (1975), "The Fishbein Extended Model and Consumer Behavior," Journal of Consumer Research, 2, September, 118-36.

Ryan, Michael J. and E. H. Bonfield (1977), "Fishbein's Extended Model: A Test of External and Pragmatic Validity," New York: Columbia University Graduate School of Business, Research Paper Number 168.

Ryan, Michael J. and Michael J. Etzel (1976), "The Nature of Salient Outcomes and Referents in the Extended Model," in Advances in Consumer Research, Volume III, ed. Beverlee B. Anderson, Association for Consumer Research, 485-90.

Sheth, Jagdish N. (1969), "Using Factor Analysis to Estimate Parameters," Journal of the American Statistical Association, 64, September, 808-22.

Stanton, John L. and Jeffrey A. Lowenhar (1974), "A Congruence Model of Brand Preference: A Theoretical and Empirical Study," Journal of Marketing Research, 11, November, 427-33.

Wilkie, William L. and Rolf P. Weinreich (1972), "Effects of the Number and Type of Attributes Included in an Attitude Model: More Is not Better," in Proceedings of the Third Annual Conference of the Association for Consumer Research, ed. M. Venkatesan, 325-40.

Wilkie, William L. and Edgar A. Pessemier (1973), "Issues in Marketing's Use of Multi-Attribute Attitude Models," Journal of Marketing Research, 10, November, 428-41.

Zaltman, Gerald, Christian R. A. Pinson, and Reinhard Angelmar (1973), Metatheory and Consumer Research, New York: Holt, Rinehart and Winston.

## Conclusion

While starting on a highly critical note, this paper ends more hopefully. It should be clear that multi-attribute models of attitude can be developed which are valid, which provide an understanding of attitude, and which lead to strategies for changing attitude. At the same time, it should be clear that attitudes alone do not control behavior.

Naturally, all problems associated with multi-attribute models cannot be discussed within the limited scope of this paper. Such important topics as consistency of attribute meaning (Gensch and Golob, 1975), attribute content (Ahtola, 1975; Etter, 1975), and specific combination rules, which include the question of weighting, have not been addressed.

Finally, a more global multi-attribute model which is not a model of attitude *per se* may lead to a parsimonious model for prediction, influence, and understanding of consumer behavior.

## References

- Ahtola, Olli T. (1975), "The Vector Model of Preferences: An Alternative to the Fishbein Model," Journal of Marketing Research, 12, February, 52-9.
- Arndt, Johan and Edgar Crane (1975), "Response Bias, Yea-Saying, and the Double Negative," Journal of Marketing Research, 12, May, 218-20.
- Baker, Michael J. and Gilbert A. Churchill, Jr. (1977), "The Impact of Physically Attractive Models on Advertising Evaluations," Journal of Marketing Research, 14, November, 538-55.
- Barnes, James G. (1978), "A Hierarchical Model of Source Effect in Retail Newspaper Advertising," in Advances in Consumer Research, Volume V, ed. H. Keith Hunt, Association for Consumer Research, 235-42.
- Bauer, Raymond A. (1966), "Attitudes, Verbal Behavior, and Other Behavior," in Attitude Research at Sea, ed. Lee Adler and Irving Crespi, Chicago: American Marketing Association, 3-14.
- Bernardo, John J. and J. M. Blin (1977), "A Programming Model of Consumer Choice Among Multi-Attributed Brands," Journal of Consumer Research, 4, September, 111-9.
- Bither, Stewart and Peter Wright (1977), "Preference Between Product Consultants: Choices vs. Preference Functions," Journal of Consumer Research, 4, June, 39-47.
- Bonfield, E. H. (1974), "Attitude, Social Influence, Personal Norm, and Intention Interactions as Related to Brand Purchase Behavior," Journal of Marketing Research, 11, November, 379-89.
- Cohen, Joel B. (1974), "Toward an Integrated Use of Expectancy-Value Attitude Models," in Buyer/Consumer Information Processing, ed. G. David Hughes and Michael L. Ray, Chapel Hill: University of North Carolina Press, 331-46.
- Cohen, Joel B. and Olli T. Ahtola (1971), "An Expectancy X Value Analysis of the Relationship Between Consumer Attitudes and Behavior," Proceedings of the 2nd Annual Conference of the Association for Consumer Research, ed. David M. Gardner, 344-64.
- Cox, Eli P., III (1975), "Family Purchase Decision Making and the Process of Adjustment," Journal of Marketing Research, 12, May, 189-95.
- Doob, Leonard W. (1947), "The Behavior of Attitudes," Psychological Review, 54, 135-56.
- Etter, William L. (1975), "Attitude Theory and Decision Theory: Where is the Common Ground?" Journal of Marketing Research, 12, 481-3.
- Fishbein, Martin (1967a), "A Behavior Theory Approach to the Relations Between Beliefs about an Object and the Attitude Toward the Object," in Readings in Attitude Theory and Measurement, ed. Martin Fishbein, New York: Wiley, 389-400.
- Fishbein, Martin (1967b), "Attitude and the Prediction of Behavior," in Readings in Attitude Theory and Measurement, ed. Martin Fishbein, New York: Wiley, 477-92.
- Fishbein, Martin and Icek Ajzen (1975), Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research, Reading, Mass.: Addison-Wesley.
- Gensch, Dennis H. and Thomas F. Golob (1975), "Testing the Consistency of Attribute Meaning in Empirical Concept Testing," Journal of Marketing Research, 12, August, 348-54.
- Hall, Calvin S. and Gardner Lindzey (1968), "The Relevance of Freudian Psychology and Related Viewpoints for the Social Sciences," in The Handbook of Social Psychology, Second Edition, ed. Gardner Lindzey and Elliot Aronson, Reading, Mass.: Addison-Wesley, Volume One, 245-319.
- Hawkins, Del I., Gerald Albaum, and Roger Best (1974), "Stapel Scale or Semantic Differential in Marketing Research," Journal of Marketing Research, 11, August, 318-22.
- Holbrook, Morris B. (1977), "Comparing Multiattribute Attitude Models by Optimal Scaling," Journal of Consumer Research, 4, December, 165-71.
- Howard, John A. and Jagdish N. Sheth (1969), The Theory of Buyer Behavior, New York: Wiley.
- Kassarjian, Harold H. (1978), "Presidential Address, 1977: Anthromorphism and Parsimony," in Advances in Consumer Research, Volume V, ed. H. Keith Hunt, Association for Consumer Research, xiii-iv.
- Kassarjian, Harold H. and Waltrud M. Kassarjian (1978), "Attitude Under Low Involvement Conditions," in Attitude Research Plays for High Stakes, ed. John C. Maloney and Bernard Silverman, Chicago: American Marketing Association, forthcoming.
- Lastovicka, John L. (1978), "Are Attitude Models Appropriate for Mass TV Advertising?" Paper presented at the American Marketing Association Ninth Annual Attitude Research Conference, Tarpon Springs, Fla.
- LaPiere, Richard T. (1934), "Attitudes versus Actions," Social Forces, 13, 230-7.
- Lundstrom, William J. and Lawrence M. Lamont (1976), "The Development of a Scale to Measure Consumer Discontent," Journal of Marketing Research, 13, November, 373-81.
- Lutz, Richard J. (1972), "Investigating the Feasibility of Personalized Rapid Transit: An Experimental Approach," in Proceedings of the Third Annual Conference of the Association for Consumer Research, ed. M. Venkatesan, 800-6.
- Lutz, Richard L. (1975), "Changing Brand Attitudes Through Modification of Cognitive Structure," Journal of

$$A^* = u_x (2P_x - 1)/k + u_y (2P_y - 1)/k \quad (7)$$

or

$$A^* = a_x b_x + a_y b_y = A_o \quad (8)$$

if

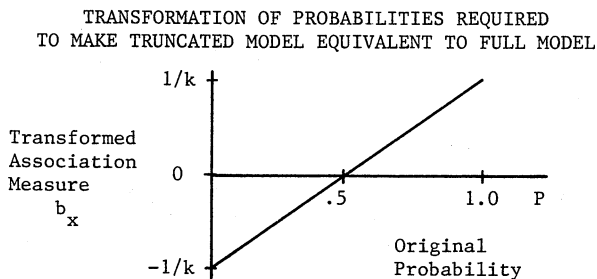
$$a_x = u_x, a_y = u_y \quad (9)$$

and

$$b_x = (2P_x - 1)/k, b_y = (2P_y - 1)/k. \quad (10)$$

The above analysis indicates that the two models produce proportional results (up to a multiplicative constant,  $k$ ), if the Fishbein measure of affect equals the utility score, and if the Fishbein measures of association are derived from a symmetric rescaling of the marginal probabilities. The form of this rescaling is illustrated in Figure 2.

FIGURE 2



Notice that this recoding does not imply that probabilities actually range from  $-1/k$  to  $1/k$ . On the contrary, it asserts that the probabilities must be modified to satisfy the axioms of probability given a truncated model. This result is consistent with Fishbein's assertion that a probability of less than 50% of achieving a desired attribute results in a lessening of the total attitude score. In the full model this lessening occurs when  $P_x$  is less than 50% because the negation of a desired attribute ( $-u_x$ ) times its occurrence ( $1-P_x$ ) is larger than the value of the positive occurrence ( $u_x$ ) times its occurrence ( $P_x$ ). In the truncated model the coding of the association measure to be symmetrical about zero is necessary to produce this same effect.

Notice further that this result, while framed in the context of the Fishbein Model, applies to all expected utility models. That is, if a researcher uses bipolar attributes and includes only one pole in the evaluation of an object, then the probabilities should be recoded to be symmetric about zero to account for the pole that was left out. This result, however, depends critically on the validity of the assumptions which are discussed in the next section.

#### The Validity of the Assumptions

The first two assumptions require that the subjective probabilities follow the axioms of probability. In a series of experiments Wyer (1974) provided evidence for the logical consistency of subjective probability judgments. Probabilities were measured on an 11-point scale ranging from extremely improbable (0) to extremely likely (10). While the experiments did not directly test (3) and (4), they tested more complex syllogisms that are true only if these assumptions hold. Thus these experiments provide support for the argument that subjective probabilities can be seen as being derived from mutually exclusive and exhaustive events.

It should be emphasized that these assumptions about

probabilities are fairly weak. Notice that they do not require that the probabilities be independent across attributes. For example, a car that is perceived as comfortable may be more likely to be perceived as reliable, but the use of the truncated model is not damaged by these correlated attributes. This is a somewhat surprising result, but quite fortunate since most attributes are perceptually correlated.

By contrast, the use of the truncated model does assume the independence of utilities across attributes. Thus, assumption (5) requires that the value of two components together be proportional to the sum of their values separately. It is possible to think of cases where such additivity would not hold. For example, the perceived attractiveness of the grill section of a car might interact with the perceived attractiveness of the hood, so that if one aspect is unattractive the value of the other aspect is attenuated (e.g. the Edsel). Generally, however, the utility of components can be considered additive with little loss in accuracy. For example, in the current case, the value of reliability would not be expected to change appreciably depending on the level of comfort. Furthermore, decompositional models have typically not found the deviation from additivity to be significant even in cases where such deviance was expected (Huber, 1975). Paradoxically, it is the very robustness of the additive model that has led some to question whether it provides a distorted description of the decision process (Anderson, 1971). However, from a predictive standpoint there can be little doubt that an additive model accounts for a vast majority of the variance. Thus the truncated model can be expected to be predictively quite accurate even if it fails to account for expected nonadditivity of the component values.

Of the assumptions, the bipolarity of utilities is the most problematic. It asserts that the value of attaining an attribute has the opposite sign but same value as not attaining it. It is similar to the familiar constraint in additive models that the sum of a factor's values equals zero. Since in the present case each attribute is assumed to have just two values for its sum to equal zero one must be the negation of the other.

However, it is an empirical question whether, in fact, assumption (6) is justified. Bettman, Capon and Lutz (1975), and Lutz (1976) have provided some support for the assumption in that probabilities that are coded in a bipolar fashion fit data better than those that are coded in a unipolar fashion. A more complete test requires one to estimate the full model and then see which assumptions best lead to approximation of the full model by a truncated version. It is likely that certain attributes will require different rescaling. For example, the attribute "perfect repair record" might have a very high partworth value while the attribute "not perfect repair record" could not be expected to have a correspondingly high negative value. By contrast, the value for "reliable" might reasonably be considered the negation of "not reliable." Thus it is likely that the truncated model will only work for certain kinds of attributes. For other attributes the full model or even more complex models such as Ahtola's (1975) are needed.

#### Summary

The purpose of this note has not been to test the assumptions of the model, although in the light of the inconclusiveness of testing the predictive accuracy of competing attitude models, it might be desirable for attitude researchers to focus on and test the assumptions instead of the overall predictiveness of their models. Rather, its purpose has been to argue that given certain reasonable assumptions, the Fishbein model yields

A NOTE ON THE FISHBEIN ATTITUDE MODEL  
AS AN EXPECTED UTILITY MODEL

Joel Huber, Duke University  
Robert P. Leone, The University of Texas at Austin

Abstract

Fishbein's attitude models and similar expectancy-value formulations, code the association measure as bipolar. Although there is good empirical justification for this coding, it results in the theoretically embarrassing prospect of negative probabilities. If, however, the Fishbein model is seen as a truncated version of the full expected utility model, then the bipolar measures of association are required by, rather than in contradiction of, the axioms of probability.

Introduction

One of the more appealing aspects of the Fishbein Attitude model (Fishbein and Ajzen, 1975) is its parallel with an expected utility formulation. That is, in the Fishbein model, attitude is a function of the evaluation of an attribute multiplied by its degree of association with the object. This is analogous to an expected utility model which multiplies the utility or value of each event times the probability of occurrence. This point has been made by Etter (1975, p. 482) who concluded that the Fishbein model "...can be regarded as equivalent to a multi-attribute decision theory model where the decision is made under conditions of risk and utility is additive." Unfortunately the equivalence is marred by Fishbein's bipolar scaling of association. Ahtola (1975, p. 52) has commented that "the operationalization of  $B_i$  (the measure of association) should be more in accordance with the axioms of probabilities." Other researchers have expressed similar concern with respect to this coding (Lutz, 1976; Bettman, Capon and Lutz, 1975).

The purpose of this note is to argue that the Fishbein scaling of the association measure as bipolar does not contradict the axioms of probability but follows from them. The problem lies in the fact that the Fishbein model does not include the negative outcome and so represents a truncated version of the full expected utility model. Given the full model and certain reasonable assumptions, it can be shown that the axioms of probability require that the measures of association in the truncated model be rescaled from the range 0,1 to a range that is symmetrical about zero. Thus the negative values are a mathematical necessity brought about by the use of an incomplete expected utility model.

Figure 1 shows the needed inputs to a full expected utility model for a hypothetical automobile. Here it is assumed that there are just two attributes, reliability and comfort, although the results generalize easily to any number of dimensions. Assume further that the attributes are binary, i.e., a car is either reliable ("x") or it is not ("x̄") and is either comfortable ("y") or not ("ȳ") and that an automobile may take any of the four joint states reflecting combinations of these attributes. In an expected utility model the value of a particular car is the sum of the products of the joint values times their probabilities, or

$$A^* = u_{xy} P_{xy} + u_{xȳ} P_{xȳ} + u_{x̄y} P_{x̄y} + u_{x̄ȳ} P_{x̄ȳ} \quad (1)$$

FIGURE 1

UTILITIES AND PROBABILITIES REQUIRED FOR FULL MODEL:  
TRUNCATED MODEL USES MARGINAL TOTALS

		Dimension 1 (reliability)		
		x	x̄	Total
y		$u_{xy}$ $P_{xy}$	$u_{x̄y}$ $P_{x̄y}$	$u_y$ $P_y$
	ȳ	$u_{xȳ}$ $P_{xȳ}$	$u_{x̄ȳ}$ $P_{x̄ȳ}$	$u_{ȳ}$ $P_{ȳ}$
Total		$u_x$ $P_x$	$u_{x̄}$ $P_{x̄}$	0 1

This is the full expected utility model. It can be contrasted with the simpler Fishbein model which for two attributes is:

$$A_o = a_x b_x + a_y b_y \quad (2)$$

where  $a_x$  is the marginal goodness or badness of x, and  $b_x$  is the degree of association between the object and the attribute x. As the model is operationalized, respondents are not asked questions about the value of both reliability and unreliability but are asked to respond to only one of the two poles of the attribute. The measure of association,  $b_x$ , is coded to be symmetric about zero so that high levels of association result in positive values and low levels result in negative values.

Equation (2) is certainly easier to estimate and understand than the full model. Moreover, it will be shown that given reasonable assumptions, it produces identical results to the full model if the measures of association are coded to be symmetrical about zero. Although shown for only two attributes, this result holds for any number of attributes and thus applies to the general Fishbein model.

To prove this equivalence certain assumptions are necessary about the relation of the joint to the marginal values in Figure 1. These are provided below:

<u>Assumption</u>	<u>Mathematical Form (for x, x̄, y, ȳ)</u>
Mutually Exclusive and Exhaustive Probabilities	$P_x + P_{x̄} = 1$ (3)
	$P_{xy} + P_{xȳ} = P_x$ (4)
Additivity of Utilities	$u_{xy} = (u_x + u_y)/k$ (5)
	k = any positive number
Bipolarity of Utilities	$u_x = -u_{x̄}$ (6)

If these assumptions and their comparable forms for the different arguments are inserted into Equation (1) the result is:

Arun K. Jain, State University of New York at Buffalo  
 Vijay Mahajan, Ohio State University  
 Naresh K. Malhotra (student), State University of New York at Buffalo

Abstract

This paper synthesizes the compositional and decompositional approaches to modelling consumer preferences. Towards this, the similarities and differences between the two approaches in terms of attribute identification, model formulation, data collection, parameter estimation, and reliability and validity testing are examined. Implications for modelling of consumer preferences are discussed and some extensions are proposed.

Introduction

The modelling of consumer preferences among multiattribute alternatives has been of great interest and concern to consumer researchers. Preference modelling involves identification of the choice alternatives, attributes associated with the alternatives, estimation of the relative contribution of these attributes, and the specification of a conceptual model underlying the choice process. Over the years, two different research approaches for estimating the consumer preference function have emerged: compositional and decompositional. For the compositional or build-up approach the total utility for a multiattribute alternative is obtained as a weighted sum of the alternative's perceived attribute levels and associated value ratings as separately (and explicitly) judged by the consumer. In contrast, the decompositional approach utilizes a consumer's response to a set of "total" profiles of a choice alternative to derive preference functions.

The focus of this paper is to compare and contrast the compositional and decompositional approaches. This is done by first examining the issue of identification of the relevant attributes. We then present a general model formulation which encompasses the two approaches. Next, we focus on data collection, estimation, reliability testing, and validation involved in the two approaches. The paper concludes with a brief discussion of some new developments in the field of multiattribute preference modelling.

Attribute Identification

An important, but frequently overlooked issue in the use of multiattribute decision models is the determination and selection of a relevant set of attributes. This is of concern in applications using compositional models in which respondents are asked to evaluate competing alternatives along each of several attributes, and in decompositional models in which respondents rate or rank profiles formed using various combinations of attributes. The attribute selection problem is especially critical in research employing the decompositional approach because the number of profiles to which respondents must react can become quite large if the number of attributes is not kept to a minimum. There are ways to reduce the number of profiles in the decompositional approach, given a fixed number of attributes, through the use of orthogonal arrays (Green, 1974), but the importance of keeping the number of attributes to a minimum remains. !

As noted by Wilkie and Pessemier (1973), an entirely satisfactory method for attribute generation and selection has not yet been developed. A particularly attractive

approach to attribute identification is the Kelley's Repertory grid (Kelley, 1955; Frost and Braine, 1967). It involves presenting respondents with triplets of competing brands (or choice alternatives). The respondent is asked to think of a way in which any two of the three brands are similar to each other and different from the third one. This process is continued with new triplets of competing brands, until the respondent has exhausted his repertoire of alternative bases for differentiation. The respondent is also required to provide the relative salience of each attribute he has identified in brand preference. The similarity and relative salience of the various constructs across the sample provide a basis for selecting an attribute set for use in the multiattribute choice models.

Model Formulation

The general main effects multiattribute preference model consisting of  $n$  attributes with each attribute defined at  $m_i$  levels may be formulated as:

$$A_{\ell k} = \sum_{i=1}^n \sum_{j=1}^{m_i} \alpha_{ij\ell} X_{ij\ell k} \quad (1)$$

where:

- $i = 1, 2, \dots, n$  attributes
- $j = 1, 2, \dots, m_i$  levels of the  $i$ -th attribute
- $k = 1, 2, \dots, K$  choice objects
- $\ell = 1, 2, \dots, L$  consumers

such that:

- $A_{\ell k}$  = The overall utility derived from the  $k$ -th alternative by the  $\ell$ -th consumer
- $\alpha_{ij\ell}$  = The value of the  $j$ -th level of the  $i$ -th attribute to the  $\ell$ -th consumer
- $X_{ij\ell k}$  = The level of the  $i$ -th attribute in the  $k$ -th choice alternative from the perspective of the  $\ell$ -th consumer

The formulation in (1) is a representation of the fundamental additive consumer preference model. Note, that the compositional model is a specific case of (1) when attribute levels are ignored. In such a case (1) reduces to:

$$A_{\ell k} = \sum_{i=1}^n \alpha_{i\ell} X_{i\ell k} \quad (2)$$

with:

- $\alpha_{i\ell}$  = The importance weight given to attribute  $i$  by the  $\ell$ -th consumer
- $X_{i\ell k}$  =  $\ell$ -th respondent's belief as to the extent to which attribute  $i$  is offered by choice alternative  $k$

This represents the traditional compositional model as identified by Wilkie and Pessemier (1973).

In the case of the decompositional model for consumer  $\ell$ ,  $\alpha_{ij\ell}$  is the part-worth contribution associated with the



results equivalent to the full expected utility model only if the probabilities are rescaled to be symmetrical about zero. Thus the Fishbein scaling of the association measure is not in conflict with either the axioms of probability or the expected utility model but actually follows from them.

#### References

- Oli T. Ahtola, "The Vector Model of Preferences: An Alternative to the Fishbein Model," Journal of Marketing Research, 12(February, 1975), 52-9.
- Norman Anderson, "Intregation Theory and Attitude Change," Psychological Review, 78(May, 1971), 171-206.
- J. Bettman, N. Capon and R.J. Lutz, "Cognitive Algebra in Multi-Attribute Attitude Models," Journal of Marketing Research, 12(May, 1975), 151-64.
- M. Fishbein and I. Ajzen, Belief, Attitude, Intention, and Behavior, An Introduction to Theory and Research (Reading, Mass.: Addison-Wesley, 1975).
- Joel Huber, "Predicting Preferences on Experimental Bundles of Attributes: A Comparison of Models," Journal of Marketing Research, 12(August, 1975), 290-7.
- Richard J. Lutz, "Conceptual and Operational Issues in the Extended Fishbein Model," Advances in Consumer Research, v. III, Beverlee Anderson, ed., Association for Consumer Research, 1976.
- Robert S. Wyer, Cognitive Organization and Change: An Information Processing Approach (Potamic, Maryland: Lawrence Erlbaum Associates, 1974).

- Paul E. Green, "On The Design of Choice Experiments Involving Multifactor Alternatives," Journal of Consumer Research, 1 (1974), 61-8.
- \_\_\_\_\_, Frank J. Carmone and D. P. Wachspress, "On The Analysis of Qualitative Data in Marketing Research," Journal of Marketing Research, 14 (1977), 52-9.
- \_\_\_\_\_, and Vithala R. Rao, "Conjoint Measurement for Quantifying Judgmental Data," Journal of Marketing Research, 8 (1971), 355-63.
- \_\_\_\_\_, and Donald S. Tull, Research for Marketing Decisions (Englewood Cliffs: Prentice Hall, 1978).
- \_\_\_\_\_, and Yoram Wind, Multiattribute Decisions in Marketing: A Measurement Approach (Hinsdale: Dryden Press, 1973).
- \_\_\_\_\_, and V. Srinivasan, "Conjoint Analysis in Consumer Research: Issues and Outlook," Journal of Consumer Research, Forthcoming.
- Arun K. Jain, "A Method for Investigating and Representing Implicit Theory of Social Class," Journal of Consumer Research, 2 (1975), 53-9.
- \_\_\_\_\_, Franklin Acito, Naresh K. Malhotra, and Vijay Mahajan, "A Comparison of The Internal Validity of Alternative Parameter Estimation Methods in Decomposition Multiattribute Preference Models," Working Paper, School of Management, State University of New York at Buffalo, 1978.
- Richard M. Johnson, "Tradeoff Analysis of Consumer Values," Journal of Marketing Research, 11 (1974), 121-7.
- \_\_\_\_\_, "A Simple Method for Pairwise Monotone Regression," Psychometrika, 40 (1955), 163-8.
- George A. Kelly, Psychology of Personal Constructs, Vols. I and II (New York: Norton, 1955).
- K. S. Krishnan, "Incorporating Thresholds of Indifference in Probabilistic Choice Models," Management Science, 23 (1977), 1224-33.
- Joseph B. Kruskal, "Analysis of Factorial Experiments by Estimating Monotone Transformations of The Data," Journal of The Royal Statistical Society, Series B (1965), 251-63.
- Daniel McFadden, "Conditional LOGIT Analysis of Qualitative Choice Behavior," in Zarembka, ed., Frontiers in Econometrics. (New York: Academic Press, 1970).
- Vijay Mahajan, Arun K. Jain, Arun J. Thangaraj, B. Ravichandran and Franklin Acito, "An Experiment With The Attribute-Dynamic Attitude Model (ADAM)," Proceedings, The Fourth International Research Seminar in Marketing, Senanque Abbey, Gordes, France (1977).
- Vijay Mahajan, Arun K. Jain, Arun J. Thangaraj, and Stephen A. Goodwin, "Attribute Determinance: An Evaluation of Alternative Measures and Some Potential Applications," in Shocker, ed., Proceedings, AMA/MSI/PIT Workshop on Analytical Approaches to Product and Marketing Planning, (1977).
- James H. Myers and Mark I. Alpert, "Semantic Confusion in Attitude Research: Salience vs. Importance vs. Determinance," Proceedings, Association of Consumer Research, 1976, 611-5.
- Barnett R. Parker and V. Srinivasan, "A Consumer Preference Approach to The Planning of Rural Primary Health Care Facilities," Operations Research, 24 (1976), 991-1025.
- D. Pekelman and Subrata K. Sen, "Mathematical Programming Models for The Determination of Attribute Weights," Management Science, 20 (1974), 1217-29.
- G. N. Punj and R. Staelin, "A Model of The College Choice Process," in Bernhardt, ed., Proceedings, Educator's Conference, American Marketing Association, (1976), 324-9.
- Vithala R. Rao, "Conjoint Measurement in Marketing Analysis," in Sheth, ed., Multivariate Methods for Market and Survey Research (Chicago: American Marketing Association, 1977).
- V. Srinivasan and Alan D. Shocker, "Linear Programming Techniques for Multidimensional Analysis for Preferences," Psychometrika, 38 (1973), 337-69.
- William L. Wilkie and Rolf P. Weinrich, "Effects of The Number and Types of Attributes Included in an Attitude Model: More is Not Better," Proceedings, Association for Consumer Research, (1972), 325-40.
- \_\_\_\_\_, and Edgar A. Pessemier, "Issues in Marketing's Use of Multi-Attribute Attitude Models," Journal of Marketing Research, 10 (1973), 428-41.
- M. Zeleny, "The Attribute-Dynamic Attitude Model (ADAM)," Management Science, 23 (1976), 12-26.

$j$ -th level of the  $i$ -th attribute and  $X_{ijkl}$  represents the presence (= 1) or absence (= 0) of the  $j$ -th attribute level in the  $k$ -th choice alternative.

Thus, it is observed that both the compositional and decompositional approaches can be represented by the same fundamental model (1). However, in the decompositional approach explicit recognition is accorded to the levels at which an attribute is presented in a particular choice alternative.

#### Data Collection

The operationalization of the compositional model will require the researcher to obtain information about the importance weights,  $a_{ij}$ , and beliefs,  $X_{ijkl}$ , from the consumers. This data is generally obtained on rating scales assumed to have interval properties. Considerable attention has been paid to the conceptualization, generality, measurement and halo effects associated with beliefs and importance weights. An excellent discussion of these issues is provided by the definitive work of Wilkie and Pessimier (1973) and hence is not pursued here.

In the case of the decompositional approach,  $X_{ijkl}$  are defined by the manner in which the choice set is constructed. The values of the parameters  $a_{ij}$  are estimated on the basis of the reactions of the respondents to the choice set. The reaction of the respondent to the choice set is generally obtained through one of two approaches: A. full profile and B. trade off.

In the full profile approach respondents are presented with a set of choice alternatives. These alternatives are described in terms of all associated attributes. The attribute levels across the choice set are varied systematically according to some experimental design (Green, 1974). The respondent is required to rank order

the alternatives in the choice set in order of his preference. On the other hand in the trade off approach the respondent is asked to rank the various combinations of each pair of attribute levels from most preferred to least preferred. Figure 1 shows an illustration of these two approaches as applied to consumer evaluations of banks.

Thus, it will be observed that the respondent is asked to provide a greater amount of information in the compositional approach as compared to the decompositional approach. Furthermore, while in the decompositional approach the respondent is required to provide only ordinal (or weaker, e.g., categorical) information, the compositional approach requires the respondent to provide the data on an interval scale. However, it should be noted that in the decompositional approach the respondent is required to process a greater amount of information since he is forced to make trade offs between alternatives defined in terms of two or more attributes.

#### Estimation

The decompositional approach requires the use of some optimizing procedure to estimate the model parameters,  $a_{ij}$ . This is not essential in the case of the compositional approach. The major focus of current research in model estimation under the two approaches is highlighted in the following.

#### Compositional Approach

An important concern among the researchers using compositional approach has been to represent attribute determinance in model formulation. Myers and Alpert (1976) suggest that attribute determinance includes but goes beyond "importance." They argue that it is the attribute determinance which should be included in the fundamental

FIGURE 1

### ALTERNATIVE DATA COLLECTION METHODS IN DECOMPOSITION MULTI-ATTRIBUTE PREFERENCE MODELS

#### I. Trade-Off Approach (Example of Two Checking-Account-Attributes-At-A-Time)

##### COST OF MAINTAINING CHECKING ACCOUNT

ACCESSIBILITY TO BANKING SERVICE	Each check you write will cost 15¢	A minimum balance of \$200 is required <u>all</u> the time in your savings or checking account. A monthly charge of \$2.00 is levied for any balance below \$200.	<u>No</u> minimum balance in checking or savings account required (that is, absolutely free)
Banking facilities are available within a <u>5</u> minute drive from <u>your</u> home.			
Banking facilities are available within a <u>10</u> minute drive from <u>your</u> home.			
Banking facilities are available within a <u>15</u> minute drive from <u>your</u> home.			

FIGURE 1 (CONTINUED)

II. Full-Profile Approach  
(Example of Checking Account Full-Profile)

Cost:	Each check you write will cost 15¢.
Type:	Commercial bank with headquarters in City A.
Accessibility:	Banking facilities are available within a 10 minute drive from your home.
Service:	Service is less friendly than average. Bank personnel are less likely to go out of their way to help you compared to most banks.
Hours:	Weekdays 9:00 AM - 4:00 PM and Evenings 5:00 PM - 8:00 PM twice a week, and Saturdays 9:00 AM - 12:00 Noon

model (1).

An approach for operationalizing determinance involved (Alpert, 1971):

$$D_{i\ell} = \lambda_{i\ell} \alpha_{i\ell}$$

where:

$D_{i\ell}$  = Determinance of the  $i$ -th attribute for the  $\ell$ -th consumer

$\lambda_{i\ell}$  = Perceived differences among choices in the  $i$ -th attribute by the  $\ell$ -th consumer

In light of the above, model (2) can be reformulated as follows:

$$A_{\ell k} = \sum_{i=1}^n D_{i\ell} X_{ik\ell} \quad (3)$$

The various approaches to identify  $\lambda_{i\ell}$  may be broadly classified as: (a) direct questioning, (b) indirect questioning, and (c) observation and experimentation. In the direct questioning approach, the respondents may be asked to indicate perceived differences among the choice set on the preselected attributes. The various indirect approaches to identify  $\lambda_{i\ell}$  include covariate analysis (Alpert, 1971; Myers and Alpert, 1976), standard deviation (Berkowitz, et al., 1976; Mahajan, Jain, Thangaraj, Ravichandran and Acito, 1977; Wilkie and Weinrich, 1972), and entropy (Mahajan, Jain, Thangaraj, Ravichandran and Acito, 1977; Wilkie and Pessemier, 1973). The third approach -- observation and experimentation -- would involve varying the size of the choice set, the number of attributes, and their levels. The effect of such variations should be reflected in the values of  $\lambda_{i\ell}$ . In particular, indirect approaches to identify  $\lambda_{i\ell}$  have been in favor with researchers in consumer behavior (Mahajan, Jain, Thangaraj and Goodwin, 1977). For example, Wilkie and Weinrich (1972) have used standard deviation as a measure of  $\lambda_{i\ell}$ . More specifically, they defined the determinance of an attribute  $i$ :

$$D_{i\ell} = \sigma_{i\ell} \alpha_{i\ell}^* \quad (4)$$

where, the importance weight of the  $i$ -th attribute is standardized across the choice set ( $\alpha_{i\ell}^* = \alpha_{i\ell} / \sum_{i=1}^n \alpha_{i\ell}$ ). Berkowitz, et al., (1976) have employed the standardized

standard deviation,  $\sigma_{i\ell}^* = \sigma_{i\ell} / \sum_{i=1}^n \sigma_{i\ell}$ , in defining the determinance of attribute  $i$ :

$$D_{i\ell} = \sigma_{i\ell}^* \alpha_{i\ell}^* \quad (5)$$

Zeleny (1976) prefers to label  $\lambda_{i\ell}$  as contrast intensity among the choices in the  $i$ -th attribute rather than perceived differences as suggested by Alpert (1971). Furthermore, Zeleny has labeled  $D_i$ , the determinance attribute, as the dynamic attribute weight. Finally, Zeleny labels stated importance weight,  $\alpha_{i\ell}$ , as static attribute weight. He redefines (3) as follows:

$$A_{k\ell} = \sum_{i=1}^n D_{i\ell} F_{ik\ell} \quad (6)$$

where:

$$F_{ik\ell} = X_{ik\ell} / X_{ik\ell}^*$$

In (6),  $X_{ik\ell}^*$ , labeled as anchor point, represents the maximum value on each attribute across the choice set for the  $\ell$ -th respondent ( $X_{ik\ell}^* = \max_k (X_{ik\ell})$ ), and  $F_{ik\ell}$  describes the relative proximity of a particular alternative to the respondent's anchor point. Mahajan, Jain, Thangaraj and Goodwin (1977) have proposed the use of standardized standard deviation, based on  $F_{ik\ell}$ , as a measure of contrast intensity.

Decompositional Approach

The various methods to estimate parameters in the decompositional approach may be broadly classified as (Jain, et al., 1978):

- Monotone regression methods such as MONANOVA (Carmone, et al., 1978; Green and Rao, 1971; Green, Wind and Jain, 1972; Green and Wind, 1973; Green, et al., 1977; Green and Tull, 1978; Jain, 1975; Kruskal, 1965; and Rao, 1977), and JOHNSON (Davidson, 1973; Johnson, 1974, 1975);
- Mathematical programming methods such as LINMAP (Parker and Srinivasan, 1976; Pekelman and Sen, 1974; Srinivasan and Shocker, 1973);
- Econometric methods such as Ordinary Least Squares (OLS); and,
- Stochastic modelling methods such as the LOGIT, PROBIT and TOBIT models (Cox, 1977; Doyle, 1977; Krishnan, 1977; MacFadden, 1970; Punj and Staelin, 1976).

The common objective of all these procedures is to derive interval scaled partworths,  $\alpha_{ij\ell}$ , from ordinal responses. The monotone regression methods utilize gradient-type search techniques to minimize iteratively a measure of badness of fit (referred to as "theta" or "stress") such that the predicted rankings of the choice alternatives reproduce, as closely as possible, the rankings provided by the respondent. Mathematical programming methods use techniques such as linear programming which attempt to minimize the number of violations in terms of the recovery of respondents' preferences. As compared to monotone regression methods, mathematical programming methods permit constraints on partworths and guarantee the global optimum (although there may be more than one solution). The ordinary least squares methods minimize sum of squared deviations between the observed and predicted preference values. On the other hand, stochastic models such as LOGIT model maximize a likelihood function to estimate partworths. Both OLS and LOGIT provide standard errors of partworths in addition to the global optimum. A detailed empirical comparison of these estimation procedures is provided by Jain, et al., (1978).

Olli T. Ahtola, University of Florida

## Abstract

The three papers in this session attempt to clarify some important issues relevant to the study of attitudes. With minor exceptions I would call the points of view taken to be mainly methodological. Very little is said about underlying theoretical notions.

Bringing in the basic theories would have been very useful, because the papers attempt to relate different approaches to attitude measurement to each other, which cannot be accomplished adequately by only pointing out similarities or differences in resultant mathematical forms or operational measures and methods chosen. Everybody agrees that some differences exist, but it remains still unclear whether the differences are merely operational issues or based on differences in underlying theories. Bonfield seems to take one extreme position, that is, if you do not operationalize and measure attitudes identically to "traditional" methods or do not demonstrate high correlation with a "traditional" measure, you should not call the concept you are measuring "attitude". At the other extreme, the authors of the other two papers seem to feel that many different models are not really different but only incomplete or truncated versions of more complete models.

## The Bonfield Paper

## Contributions of the Paper

Even though we all know it, it is useful to be reminded how important it is to validate our models and measuring instruments carefully and thoroughly, and to receive suggestions how to do it for attitude models. It is also good to be reminded that an attitude is a construct which need not to be directly related to overt behaviors. Furthermore, the often made assumption for methodological convenience, i.e., that the same attributes are salient for all subjects in our population and for all alternatives in a choice set, may not be a good assumption to make in many cases. It is also a good suggestion that for behavior prediction perhaps other models than multiattribute attitude models, might be more useful.

## Questions with the Paper

I feel the author is taking a very extreme approach when he argues that the Thurstone equal appearing intervals attitude scale is a measure of attitudes by definition and that all other methods must correlate with it in order to be valid measures of attitudes. Personally, I strongly disagree that any method to measure a construct, especially a hypothetical construct is valid by definition. I think, everybody agrees that at least one of the dimensions of attitude is affect, but the fact that Thurstone was perhaps the first who proposed and tried to justify a method to measure affect toward objects on an interval scale, does not give this method any seniority-based right to claim ownership of the concept it was intended to measure.

Of course, it is sensible to find out if a new method converges with the already established methods as a part of the validation process, especially if the new method is based on the same conceptualization of the construct as the established methods. However, I do not feel this comparison is absolutely necessary, nor do I feel it is adequate. To the best of my knowledge,

the Guttman scales, for example, became quite established before anybody correlated them with the Thurstone scales. What makes a method valid, including the Thurstone scales, is that it follows logically from the underlying theory, predicts what it is theoretically supposed to predict and does not covary with theoretically unrelated variables. (Campbell and Fiske, 1959; Cronbach and Meehl, 1955)

It is better to consider some methods to measure attitudes to be less (or poorly, or perhaps incorrectly) validated than some other methods, than to categorize them into the dichotomy "validated/unvalidated" based on some weak criterion such as the one suggested by the author. Also, if the developers of these "unvalidated" methods would be required to label the resultant measures as measures of something that has not been measured before, as the author suggests, the result would be an utter chaos. Such labels as "values", "beliefs", "preferences" or "intentions" would not do because methods to measure them have already been proposed.

The rest of the discussion under the subtitle "A Definition of Attitude" is devoted to the reasons why attitudes do not necessarily predict overt behaviors. During this lengthy discussion, the reader is kept wondering what this issue has to do with attitude measurement. Finally, in the last sentence the reader is given a clue. The author is probably trying to tell that even if the "attitude" measure does not predict a certain overt behavior it is not necessarily invalid nor is the "attitude" measure necessarily valid if it happens to predict the behavior quite well. I think the author is quite right here if I interpret him correctly, but this part of the paper is unclearly written. During this discussion the author also, for some reason, tries to draw parallels between attitude and id. I think this is quite unnecessary and probably wrong. Quoting the same authors (Hall and Lindzey, 1968) id is conceptualized as the first, obscure, inaccessible, and unorganized part of personality, with no direct relations with external world, and with immortal contents. These characteristics certainly do not seem consistent with the typical conceptualization of attitude. Those who are interested in the psychoanalytic approach to attitudes should find Sarnoff (1960) interesting.

The author suggests the use of a properly constructed evaluative dimension of the Semantic Differential as the validation measure for new attitude scales. I think, because of its relative simplicity it is a sensible part of the validation process. However, as I already stated much more is needed to establish strong validity. Also, it is so well established that scales like "good-bad", "pleasant-unpleasant", "satisfactory-dissatisfactory" always load highly on the evaluative dimension, that if they are used without factor analytic validation, no big harm is committed. Actually, if those scales would not all have high loadings on the same factor, I would not accept that factor as the evaluative dimension of the Semantic Differential.

As I questioned the need to bring psychoanalytic notions to the discussion, I am questioning the usefulness of bringing in Field Theory to complicate the issue. Field Theory is a comprehensive theory of motivation and it is not very clear how to relate its components to attitudes. None of the multiattribute attitude theories (e.g., Fishbein, Rosenberg) state that same attributes are used by all people for all alternatives in a choice set.

Also Field Theory is not needed to justify the inclusion of "nonevaluative" attributes. Fishbein (1967, p. 396) states that any attribute (belief) associated with the attitude object can be a determinant of the attitude. However, if we get out of the realm of attitudes per se, and move toward behavior prediction then perhaps Field Theory provides more direct guidance.

In summary, even though I feel this paper correctly points out deficiencies in attitude measurement in consumer behavior, I am not equally enthusiastic about most of its suggestions. Also, I feel, unnecessary theoretical confusion was added by bringing into discussion in an ad hoc fashion some basic notions from psychoanalytic and Field Theory.

#### The Huber and Leone Paper

##### Contributions of the Paper

The authors provide reasons why the Fishbein model often provides fairly accurate predictions even though it is not conceptually correct. They also provide a model which is more sound, even though, in my opinion, it is still not quite conceptually correct either.

##### Questions with the Paper

The authors are incorrect in stating that the Fishbein model does not include negative outcomes. Any belief statement about the attitude object is acceptable to the model whether it deals with positive or negative outcomes. Perhaps the authors actually mean that after an outcome has been determined, its opposite (complement) is not incorporated in the model even though it in fact belongs there. However, Fishbein manages to salvage the model by measuring probabilities in bipolar fashion, because the "utility/disutility" of this opposite (complement) consequence is (-1) times the "utility/disutility" of the included consequence. Even though, personally I partly agree with the authors' explanation why the model works, I am fairly confident Fishbein would not. I am confident he does not see the bipolarity of belief strengths as a methodological shortcut to truncate the true underlying model. I think he would argue that a salient disassociation from something good is bad irrespective of the value of its opposite and that a disassociation from something bad is good irrespective of the value of its opposite.

My personal problem with the paper is that I do not believe the assumption of "bipolarity of utilities" is in many cases realistic, nor do I feel it is adequate to include only a salient outcome and its opposite in the model. Most outcomes are not perceived to be simply dichotomous, but in fact have several levels, of which all salient ones must be incorporated into the model.

Why make the unreasonable assumptions of dichotomy of attribute levels and bipolarity of utilities if it is not necessary. Ahtola (1975) has proposed the model which does not rely on these assumptions. A good example where a dichotomous salient attribute does not meet the required "bipolarity of utilities" is sex. Sex of an insurance salesperson may be a salient attribute to a client. He may feel that female is OK but male would be even better or vice versa. Thus, both attribute categories have positive or at least non-negative "utility". Such attribute as "temperature of tea" is perceived to have several levels and it is certainly unlikely that most people would evaluate the extreme categories (hot/cold) to have opposite evaluations.

In summary, if the purpose of this paper is to suggest a reason why the Fishbein model sometimes works in spite of being conceptually incorrect, I agree with the authors, but if the purpose is to justify the conceptualization of the Fishbein model, I disagree with the authors. It is not clear which stand the authors are taking.

#### The Jain, Mahajan, and Malhotra Paper

##### Contributions of the Paper

This paper is especially useful for someone interested in multiattribute preference modeling, but not willing to go through the trouble of reading the original literature. An amazing amount of literature is described and compared in only four pages.

##### Questions with the Paper

By attempting to do so much in only four pages it is natural that several important issues are touched very lightly or totally left out. Also, perhaps for the same reason, the paper is very descriptive with little critical analyzing.

Somewhat surprisingly, one of the issues totally left out is the theoretical justification for the model(s) and the implications of the theory/theories to the issues discussed. Also, the well known and closely related "Expectancy-Value" type of multiattribute preference models (see, for example, Cohen, Fishbein and Ahtola, 1972) have not at all been discussed in the paper. It is unfortunate because these models have stronger theoretical bases than the models described in the paper.

At the end of the discussion of attribute identification the authors recommend the use of Kelley's Repertory grid type of method. At the same time, they also emphasize the need to keep the number of attributes to the minimum. One problem with the grid method is that it tends to make previously nonsalient attributes salient when respondents struggle to find differences where they otherwise would not perceive them. This, of course, is a threat to the external validity and adds to the number of attributes. I personally prefer a free elicitation method (Ahtola, 1973).

It might be of interest to note that the "fundamental additive model" proposed by the authors is identical to Ahtola's Vector model (Ahtola, 1975) with only one difference. This crucial difference is that I defined  $X_{ijkl}$  as the  $l$ th consumer's subjective probability that  $k$ th alternative possesses attribute  $i$  at the level  $j$ . My model was developed to solve some conceptual problems I felt were present in the Fishbein model.

Going back to the models discussed in the paper, I find it difficult to see the close connection between the "general model" and the "traditional compositional model" and for that reason the connection between the commonly used compositional and decompositional models. In the "fundamental model"  $a_{nj1}$  is a measure of value while in the "traditional compositional model" it is a measure of attribute importance. These two concepts are not identical (Cohen, Fishbein and Ahtola, 1972).

Also, the "traditional compositional models" vary widely in how  $X_{ilk}$  is conceptualized. Very commonly, it is defined in terms of how satisfactory the  $l$ th consumer feels alternative  $k$  is on attribute  $i$ . That is,  $X_{ilk}$  is often a measure of evaluation in the traditional model while in the fundamental (and compositional) model it is a measure of perception.

Perhaps in their eagerness to show that the commonly used compositional and decompositional multiattribute preference models are based on the same underlying "fundamental model" the authors have not recognized some fundamental differences among the compositional models and between these models and the proposed "fundamental model".

The only question I have about decompositional models is why so much energy is spent to derive interval scaled measures from ordinal responses. Why not measure the responses on interval scales in the first place?

#### Directions for Further Research

Research using or developing multiattribute attitude models has suffered from lack of theoretical justification, from lack of proper validation, as Bonfield so forcefully demonstrated, and from the bewildering number of slightly different conceptualizations and operationalizations. Very little conceptual and/or empirical work has been done to help the potential users of these models to select the most valid and useful one for their purposes. To some extent Jain, Mahajan, and Malhotra attempted to provide this sort of assistance. Much more work needs to be done, however. The comparative studies need to be sensitive to real differences among the models, describe the theoretical bases of the models, evaluate and empirically test the reliability and validity of alternative operationalizations of the same model, and demonstrate conceptually and empirically why a given approach is superior in a given situation, taking into consideration the purpose of the research.

#### References

- Olli T. Ahtola, An Investigation of Cognitive Structure within Expectancy-Value Response Models (Doctoral Dissertation, University of Illinois) Ann Arbor, Michigan, University Microfilms, 1973, No. 74-5517.
- Olli T. Ahtola, "The Vector Model of Preferences: An Alternative to the Fishbein Model," Journal of Marketing Research, 12 (February 1975) 52-9.
- Donald T. Campbell and Donald W. Fiske, "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix," Psychological Bulletin, 56 (1959), 81-105.
- 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), 456-60.
- Lee J. Cronbach and Paul E. Meehl, "Construct Validity in Psychological Tests," Psychological Bulletin, 52 (1955), 281-302.
- Martin Fishbein, "A Behavior Theory Approach to the Relations between Beliefs about an Object and the Attitude toward the Object," in Readings in Attitude Theory and Measurement, M. Fishbein (ed.), (New York: John Wiley & Sons, 1967), 389-400.
- Calvin S. Hall and Gardner Lindzey, "The Relevance of Freudian Psychology and Related Viewpoints for the Social Sciences," in The Handbook of Social Psychology, 2nd Edition, G. Lindzey and E. Aronson (eds.), (Reading, Mass.: Addison - Wesley, 1968), Volume One, 245-319.
- Irving Sarnoff, "Psychoanalytic Theory and Social Attitudes," Public Opinion Quarterly, 24 (1960), 251-79.

## ATTITUDE MODELS IN CONSUMER RESEARCH

Sadaomi Oshikawa, University of Washington

### Abstract

The subject of attitude models has been the topic of many research papers. Much confusion exists about the nature of the Rosenberg, Fishbein, and multi-attribute models and about how they should be operationalized to measure attitudes. The three papers which have been assigned to me have attempted to synthesize different approaches to modeling consumer preferences, to put the Fishbein model within the framework of an expected utility model, and to point out the researchers' failure to validate the attitude scales they have developed. I believe all of the three papers have contributed in their unique ways to the further understanding of attitude models.

In the following sections, each of the papers will be critically evaluated and their contributions and weaknesses pointed out.

Finally, in the "Directions for Future Research" section, the importance of establishing predictive validity of the models will be emphasized and one hypothesis about why multi-attribute models produce high correlations will be explicated.

### The Jain, Mahajan and Malhotra Paper

#### Contributions of the Paper

Jain, Mahajan and Malhotra have done an admirable job of reviewing the literature on multi-attribute models and comparing the differences in the compositional and decompositional approaches. This review article is a worthwhile reading for those who wish to keep themselves up-to-date on the topic of multiattribute models.

#### Questions with the Paper

While multiattribute models have been accepted by many consumer behavior researchers, their construct validity has never been clearly established. While Fishbein and Rosenberg attitude models have theoretical bases in stimulus-response learning theory and cognitive instrumentality theory respectively, multiattribute models in the form presented in the paper do not seem to have any theoretical base in social psychology. Research efforts need to be directed at validating multiattribute models and explaining why importance weights and respondents' beliefs are related to consumer attitudes.

The authors are interested in the validity of multi-attribute models and suggest obtaining the correlation between the attitude scores and the input data such as independent ratings of preference ordering of the alternatives. Methodological problems of this procedure will be discussed in the "Directions for Future Research" section.

### The Huber and Leone Paper

#### Contributions of the Paper

Huber and Leone have addressed themselves to the issue of unipolar and bipolar coding in attitude scales. This issue is important because there is no consensus about which coding scheme to adopt and because the choice of

unipolar or bipolar coding makes a difference in the prediction of preference among a set of attitude objects. A simple algebra can show that predicted preference of attitude objects can be reversed by using different coding schemes.

Huber and Leone are to be commended for attempting to resolve the unipolarity-bipolarity coding issue. They hypothesize that Fishbein attitude model is a truncated version of the full expected utility model and that the bipolar coding scheme for association is required by the axioms of probability. If their hypotheses and the underlying assumptions are supported, Huber and Leone will be making a great contribution to the field of consumer research.

The present writer agrees with them when they emphasize the importance of testing the assumptions underlying different attitude models instead of comparing their overall predictive capabilities.

#### Questions with the Paper

As was indicated, Fishbein developed his original model on the basis of S-R learning theory, and the attitude toward an object is the affect (or liking) toward that object. Since the affect in the original Fishbein model is both conscious and unconscious processes, it is difficult to regard the original Fishbein model as equivalent to highly cognitive expected utility model. Fishbein's attitude-toward-act model, on the other hand, was developed from Dulany's cognitive theory of propositional control, and Huber and Leone are quite right in stating that Fishbein's attitude-toward-act model is analogous to an expected utility model.

However, the authors appear to be in error in stating that the Fishbein model does not include the negative outcome because there is no theoretical reason why the negative outcome cannot be included (Fishbein, 1967, 257-66) and Fishbein and Ajzen (1975, 223) indicate in the footnote that "the model deals only with associative relations between object and attribute, and thus a belief such as 'O is not X' is viewed as an association between the object O and the attribute not X."

If the argument that the Fishbein model is a truncated version of expected utility model is based solely on the assumption that the Fishbein model does not include the negative outcome, then Huber and Leone's argument does not seem to be tenable. However, it is still possible that the Fishbein model is a truncated version of the expected utility model for the reasons they have not considered. If that is the case, the authors have contributed considerably to the field of consumer research.

Of the assumptions which must be satisfied, Huber and Leone indicate that the truncated model assumes the independence of utilities across attributes. In consumer research, halo effects are common. Product attributes interact and their utilities are not independent of each other.

Finally, the authors suggest that whether the assumption of the bipolarity of utilities is justified is an empirical question. This can be tested by examining whether probabilities which are coded bipolarly fit



the data better than those which are coded unipolarly. However, problem arises when some researchers find the bipolar coding fits the data better (Bettman, Capon and Lutz, 1975; Lutz, 1977) while others find the unipolar coding fits the data better (Ahtola, 1975).

It may be more fruitful to examine whether the assumption of the bipolarity of utilities is justified theoretically and then test it empirically. Huber and Leone have done an excellent job in directing our attention to the right (theoretical) direction.

#### The Bonfield Paper

##### Contributions of the Paper

While Jacob Jacoby (1975, 1978) had made the same point, Bonfield made a contribution by pointing out the fact that many researchers have developed but have not validated their own attitude measuring instruments. When the instruments have not been validated, the researchers' empirical results cannot be conclusive.

If the results fail to support the model, the researcher must decide whether the model is not tenable or, while the model is tenable, the measurement procedure is faulty. If the results support the model, on the other hand, he still has to make sure that the results are not due to the artifacts of design or some spurious relationships.

##### Questions with the Paper

The author states that an attitude measuring instrument can be validated by showing high correlations with Thurstone scales measuring attitude toward the same object, concept, or act. It is true that the researcher who develops a new attitude instrument needs to establish that his instrument measures attitudes as do other validated instruments by means of showing high correlations with those validated instruments. In evaluating the semantic differential as a measure of attitude, for example, Osgood, Suci and Tannenbaum (1957) compared the semantic differential with Thurstone and Guttman scales. This comparison is a necessary condition but is not a sufficient condition for establishing validity because the high correlation may result if the subjects see the connection between those scales and respond consistently to both of them.

In order to establish validity, it is necessary to demonstrate that the researcher can predict correctly future behavior of the respondents on the basis of their scores on that instrument and establish predictive validity (Osgood, Suci and Tannenbaum, 1957, 104-24).

Bonfield is quite correct, on the other hand, in stating that an unvalidated attitude measure with high measure-behavior congruence would provide no understanding of the relationship between the measure and the behavior. It may be incorrect to assert the validity of an instrument solely because of the high congruence since the high congruence may result even though the instrument is not valid.

Finally, Bonfield asks whether researchers should concentrate on multi-attribute models which contain only attitudinal vectors or should develop multi-attribute models which contain not only attitudinal vectors but also other salient vectors. Since an attitude is a predisposition to evaluate an object favorably or unfavorably, only evaluative vectors which attribute good or bad qualities to that object (e.g., delicious looking, healthy, tasty, smells good, economical, hard to digest, in his example) should be included. Definitional vectors like chicken, veal scallops and \$8.25

does not contribute to his liking or disliking of this entree. According to Katz and Stotland (1959, 429), "Judgments which are purely cognitive would not fall into the category of attitudes."

When not only evaluative vectors but also other salient vectors such as definitional and social influence vectors are included, the multi-attribute model may perform better in "predicting" behavior not necessarily because the model is better but because there are more predictor variables and consequently  $R^2$  will be higher. Adjusted  $R^2$  is a more appropriate statistic to determine which model (the multi-attribute model or the global multi-attribute model) has better "predictive" power.

Almost all researchers evaluate attitude models on the basis of the coefficient of multiple correlation or some other measure of association between the criterion variable (e.g., revealed preference, ranking, etc.) and the predictor variables. This research methodology can produce misleading and untenable results, and the following section will discuss the procedure which will reduce this bias.

#### Directions for Future Research

One of the puzzles in multi-attribute attitude research is the fact that multi-attribute models perform reasonably well in "predicting" criterion attitude even though there is no apparent social psychological theory to support the models. While Fishbein and Rosenberg models are based on S-R learning theory and instrumentality theory respectively, the multi-attribute models do not have a corresponding theoretical base. A possible reason for this good performance is the tendency of the respondent with favorable brand affect to evaluate salient attributes favorably.

This cognitive consistency, also called halo effect by Beckwith and Lehmann (1975), tends to increase the correlation between the criterion variable and the predictor variables and tends to overestimate the "predictive" capability of multi-attribute models. Consistency is a valued trait, and respondents may consider the attitude measuring instrument as a test of their ability to be consistent and respond in such a way to be evaluated as intelligent (i.e., evaluation apprehension).

Since the number of attitude objects is ordinarily small (typically less than 20) and attributes are concrete rather than abstract, it is not difficult for the respondents to get the connection between attitude objects and salient attributes and respond consistently. When the correlation is reasonably high, the researcher concludes that the model's predictive performance is satisfactory or the model has predictive capability (Bruno and Wildt, 1975; Churchill, 1972; Holbrook, 1977). This usage of the word "predictive" appears to be based upon the fact that the overall affect as the criterion variable is "predicted" by predictor variables. Some writers conclude that the model has "predictive validity" when the correlation is high (Wildt and Bruno, 1974).

However, strictly speaking, predictive validity is the ability to foretell the future behavior of respondents on the basis of their scores on the scale. In almost all studies of multi-attribute models, the criterion variable has been measured immediately before or soon after the predictor variables have been measured. This procedure not only fails to test the predictive validity but also causes the respondents to become aware of the relationships between the criterion and the predictor variables and produces artifactually high correlation coefficients.

The criterion variable and the predictor variables need to be measured at two different time periods sufficiently far apart so that the respondents do not get the connection between the two variables, and irrelevant attitude objects and attributes should be included in order to disguise the true purpose of the study.

One procedure which would enable the researchers to test predictive validity is the one Osgood, Suci and Tannenbaum (1975) used to evaluate the semantic differential. This method has the advantages of preventing the respondents' attempts for consistency from distorting the correlation and of determining the directionality of causality between the overall affect and the cognitive structure regarding the salient attributes.

Finally, if the criterion and the predictor variables are measured at about the same time, the respondents are more likely to become aware of the connection between the two variables when responding to the multi-attribute scales than when responding to the Fishbein or the Rosenberg scale. As was indicated already, the attribute beliefs and importances are usually more concrete concepts while the counterparts in the Fishbein and the Rosenberg scales tend to be more abstract. The relationship between the criterion variable and the predictor variables is more apparent in the multi-attribute scales. Having rated X brand more satisfactory than Y brand on salient attributes, the respondent may express greater affect toward X brand than toward Y brand in order to appear consistent and intelligent. The relationship in the other two scales is less obvious. This artifact may explain the relatively satisfactory performance of multi-attribute models.

#### References

- O.T. Ahtola, "The Vector Model of Preferences: An Alternative to the Fishbein Model," Journal of Marketing Research, 12 (February 1975), 52-9.
- N.E. Beckwith and D.R. Lehmann, "The Importance of Halo Effects in Multi-Attribute Attitude Models," Journal of Marketing Research, 12 (August 1975) 265-75.
- J. Bettman, N.Capon and R.J. Lutz, "Cognitive Algebra in Multi-Attribute Attitude Models," Journal of Marketing Research, 12 (May 1975), 151-64.
- A.V. Bruno and A.R. Wildt, "Toward Understanding Attitude Structure: A Study of the Complementarity of Multi-Attribute Attitude Models," Journal of Consumer Research, 2 (September 1975), 137-45.
- G.A. Churchill, Jr., "Linear Attitude Models: A Study of Predictive Ability," Journal of Marketing Research, 9 (November 1972), 423-6.
- Martin Fishbein, Readings in Attitude Theory and Measurement, ed., (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)
- M.B. Holbrook, "Comparing Multiattribute Attitude Models by Optimal Scaling," Journal of Consumer Research, 4 (December 1977), 165-71.
- Jacob Jacoby, "Consumer Research: Telling It Like It Is," in Advances in Consumer Research, Vol. 3, B.A. Anderson, ed. (Atlanta: Georgia State University, 1976), 1-11.
- Jacob Jacoby, "Consumer Research: A State of the Art Review," Journal of Marketing, 42 (April 1978), 87-96.
- D. Katz and E. Stotland, "A Preliminary Statement to a Theory of Attitude Structure and Change," in Psychology: A Study of Science, Vol. 3, S. Koch (ed.), (New York: McGraw-Hill, 1959), 423-75.
- R. J. Lutz, "An Experimental Investigation of Causal Relations Among Cognitions, Affect, and Behavioral Intention," Journal of Consumer Research, 3 (March, 1977), 197-208.
- C. E. Osgood, G.J. Suci and P.H. Tannenbaum, The Measurement of Meaning (Urbana: University of Illinois Press, 1957).
- A. R. Wildt and A.V. Bruno, "The Prediction of Preference for Capital Equipment Using Linear Attitude Models," Journal of Marketing Research, 11 (May 1974), 203-5.

A COMPARISON OF TWO METHODS FOR DETERMINING  
OPTIMUM LEVELS OF PRODUCT CHARACTERISTICS

James H. Myers, University of Southern California  
Richard F. Chay, American Cyanamid, Inc.

Abstract

Optimum levels of intensity of physical attributes of food products can be determined in many ways. This paper compares results from two approaches: respondent ratings of an "ideal" product, and levels derived indirectly from cross-tabulations of beliefs ratings vs. overall evaluations. Results show that derived levels are higher than direct levels for most product characteristics.

Introduction

Product/service evaluation research is usually conducted for one or more of the following objectives: (1) diagnosis of the relative strengths and weaknesses of each product (brand) for the further objective of, (2) repositioning the firm's own offering(s), (3) redesigning existing products or brands, and (4) identifying opportunities for new products the firm might offer. Common to all of these objectives is the need to determine the particular combination of product characteristics, and the optimum levels of each characteristic, desired by the greatest number of consumers: the "ideal" product.

Determination of optimum levels for each characteristic is dependent upon the nature of the characteristics involved. Some are measured on a scale with discreet categories, others on a continuous scale. Some involve trade-offs (with price or with other features), other do not. Some involve subjective evaluations, others are objectively determined. No single research technique is available at the present time that will provide a solution for all contexts or conditions.

This study is concerned with determining optimum levels of product characteristics for an ethnic food product used for dinner occasions. The methodology required would also be applicable for most food, beverage, cosmetic, and household cleaning products; all of these types of products are composed primarily of attributes that are measured on continuous scales, are subjective, and do not involve trade-off considerations of any meaningful degrees (e.g., thickness, sweetness, saltiness, color saturation, strength of aroma, abrasiveness, carbonation). Such characteristics almost always require optimum rather than maximum levels, since consumers do not want "too much" of them. In contrast, people usually want all of the "benefits" they can get (e.g., durability, friendliness, convenience), within the normal range of human experience.

The problem is simply to determine, on a preliminary basis, optimum levels for the two to four most important product characteristics so that laboratory personnel can develop products at these levels. The matter of interaction effects is dealt with at a later stage of development, when several formulations are made up consisting of different levels of the most important characteristics that are at or near the optimum levels determined earlier.

Optimum Level Methodology

Researchers have approached the problem of determining optimum product characteristic levels in many ways:

1. Asking respondents to describe the characteristics of an ideal product using rating scales. The problems here are well known: respondents often do not know how important some features are (e.g., styling, packaging), they may be unwilling to divulge this information, they may be unable to describe an ideal product using rating scales (how would people have described a Coca Cola before this product existed?), and their ideal product ratings tend to be very similar to those for their "most preferred" brand or product (Bass, Pessemier, 1971; Carmone, 1970; Klahr, 1970; Johnson, 1971; Neidell, 1969; Wilkie, 1973).
2. Using experimental research to obtain reactions to varying combinations of product characteristics. While this is usually the best approach, it is often not feasible or even possible except in the case of food or beverage products whose formulations can be changed relatively easily. However, these changes are time consuming and expensive and they require prior information as to which 2-3 attributes are the crucial ones (Banks, 1965).
3. Using MDS and related technologies, comparing existing products directly against an ideal product, to derive ratio-scaled distance measures from which the characteristics of an ideal product can be inferred. However, the problem of clearly identifying the axes of the resulting positioning map is well known, and there is also the problem of translating the spatial location of an ideal product into measurable physical characteristics meaningful to R & D personnel (Green, 1975; Carmone, 1970). This latter problem is common to several other approaches, also.
4. Using the most preferred (for an individual) or the best selling (for the aggregate) product/brand as an estimate of the ideal product. Measurable characteristics of this product can then be used as a reference point by laboratory personnel. Of course, the best existing product of a given type might or might not be similar to an ideal product; this would not be known unless additional research were done to determine both distance and direction of the "best" product to the ideal product. (Johnson, 1971).
5. Using optimization models (e.g., LINMAP) to estimate ideal points. This is more useful when the attribute scales and anchor statements used lead to interior (i.e., on the scale) ideal points rather than exterior (i.e., beyond the range of the scale, usually at the top or highest degree--of, say, durability) (Carroll, Chang 1965; Shocker, 1974; Zufryden, 1976). Also, these approaches often require more brands/products to rate than attributes on which to rate them, just the opposite of the usual situation in product development work.
6. Inferring optimal amounts of each characteristic from cross-tabulations of characteristic ratings vs. some meaningful dependent variable (e.g., overall evaluation, usage frequency). The problem

here is that rating scale values (averages) cannot be converted directly into measurable physical characteristics, but they can be inferred from levels of these characteristics in other products rated. Also, neither this approach nor some of the others above indicate the first or second order interactions that often signal optimum combinations of product characteristics.

In view of the above shortcomings, it is obvious that there is no ideal method of determining the ideal product!

The present study involves a comparison of two of the above approaches; direct respondent ratings (#1 above) vs. derived estimates generated from cross-tabulations of each product characteristic against overall product evaluations (#6 above). The company involved was a large food processor that was interested in entering an established product category with a new product formulation.

#### Method

A total of 621 women who had used a brand in the product category in the past three months were recruited by telephone and brought to a central location facility. These women were members or friends of members of social and religious organizations in two cities: one large mid-western city and another smaller city in the southwest United States. It should be noted that women recruited from social and religious organizations are generally less representative than when conventional sampling methods are used.

Upon entering the central location facility, the women were asked to complete a background questionnaire to answer additional questions such as:

- Frequency of category usage.
- Brand used in past three months/brand used most often.
- Satisfaction with usual brand.
- Demographics.

After completing the background questionnaire, the women were asked to taste four formulations of the food product: two popular national brands, the company's own new formulation, and a regional brand relatively inferior to the other three brands tested.

Each woman was instructed to taste each product, drinking some cold water before and after the tastings. All of the women rated the four products on each of the eleven product characteristics plus an overall evaluation of the product. Pairs of words/phrases served as anchor points for a six-box scale. The characteristics selected for this taste test were developed jointly by marketing research, technical research, and marketing management departments, using focus group discussions and usage and attitude studies as input. All characteristics rated were descriptive rather than evaluative; the latter (e.g., good flavor) were eliminated since they would have had little or no diagnostic value for laboratory personnel. Examples of characteristics rated include thickness, chunkiness, saltiness, color saturation, and spiciness.

To maintain confidentiality, the actual product characteristic descriptors will not be used here, but will be designated by letters A-K. All polar descriptive words or phrases were of the form A, not A, B, not B; for example, extremely thick, not thick. To reduce the order bias associated with sequential evaluations, a strict rotation of all products and characteristics was maintained. Overall product evaluations were obtained on a six-point scale ranging from excellent to poor.

After tasting and rating the four products, the women were asked to describe their ideal products using the same attribute scales. Averages of these ratings were used as coordinates for the ideal product in attribute space.

#### Direct Ideal Points

A mean value was calculated for ideal ratings on each of the twelve product characteristics. Table 1 shows frequency distributions for each of these characteristics for all respondents combined. This exhibit

TABLE 1

DISTRIBUTION OF IDEAL RATINGS FOR PRODUCT CHARACTERISTICS: ALL RESPONDENTS COMBINED

Scale Value	A	B	C	D	E	F	G	H	I	J	K
6	18	55	126	112	248	96	109	6	202	150	76
5	221	103	140	184	239	139	191	23	223	248	181
4	257	179	153	204	119	178	194	123	149	199	259
3	173	120	96	93	14	68	75	196	35	18	75
2	50	85	70	24	1	74	28	171	9	5	26
1	42	79	36	2	-	66	24	102	3	1	4
Mean	3.7	3.5	4.1	4.4	5.2	3.9	4.3	2.7	4.9	4.8	4.3

can be interpreted as follows: If Characteristic A were thickness (i.e., extremely thick = 1, not thick = 6), the ideal product would be somewhat more than midway between extremely thick and not thick (mean = 3.7). These mean ratings show optimum levels for each characteristic as developed directly from respondent ratings.

#### Derived Ideal Points

Optimum levels for each characteristic were developed indirectly from cross-tabulations of characteristic ratings vs. overall evaluations, pooled across all products and respondents. An example is shown in Table 2 for Characteristic K (e.g., saltiness). Note how average overall evaluation ratings increase for the first four intervals on the scale, level off, and then decline. Thus, the optimum point for this attribute would be somewhere within the scale values 4 and 5. Note that Table 2 is simply a bivariate scatterplot for purposes of inspecting the form of the relationship between a dependent and an independent variable. In this case, the relationship is non-linear, with a correlation (eta coefficient) of .36.

TABLE 2

CROSS TABULATION OF RATINGS ON CHARACTERISTIC K VS OVERALL EVALUATION

Char. K Ratings	Overall Evaluation Ratings						Row Average
	1	2	3	4	5	6	
1	54	12	7	6	3	3	1.84
2	34	21	17	15	12	10	2.82
3	12	18	22	22	18	24	3.76
4	6	13	26	32	49	32	4.27
5	7	17	11	39	52	30	4.29
6	38	33	31	42	57	57	3.84

The matter of how much of Characteristic K is represented by scale values 4 and 5 cannot be determined from cross-tabulation. It can, however, be inferred from mean ratings of each test products on the same scale.

Assuming Characteristic K can be measured by suitable laboratory instruments, the desired amount of this characteristic could be determined by interpolating from the amounts in each of the products rated.

Pooling ratings across all products is necessary in this type of analysis to provide a range of levels for each of the characteristics rated. A single product, or the same level of a given characteristic for all products, provides no opportunity for the variations that are essential for locating points where overall evaluation ratings are highest (optimum). Whenever highest overall ratings are found at either extreme of the scale, it indicates that the range is not sufficient to locate the optimum point, beyond which there is too much or too little of the particular characteristic. This can be remedied either by increasing or decreasing the amount of that characteristic for one or more test products (if this is possible) or by changing the descriptors at the ends of the rating scale into more extreme statements (e.g., very salty to extremely salty). The latter approach avoids truncation and allows respondents who do not like the saltiest product to so indicate through lower overall evaluations of this product, without being grouped with people who do not think this product is too salty.

#### Segmentation

Early in the analysis it was decided to see if meaningful segments could be found, based upon overall product preferences. Since the market for this food product was dominated by two somewhat dissimilar brands, respondents were separated into three groupings based upon their overall evaluation ratings for each of these brands: (1) prefer Brand A over B, (2) prefer Brand B over A, (3) prefer both equally (i.e., rate each at the same scale interval).

This procedure produced 269 respondents who preferred Brand A by one or more scale values on the 6-point overall evaluation scale, 266 who preferred Brand B, and 86 who were neutral. Each of the first two groups was considered to be a preference segment, and each was analyzed separately. Results are shown for each of these segments separately. To make certain that the two segments identified were not simply statistical artifacts, a significance test (t) was done of the differences in overall evaluation ratings of each of the two dominant brands between the two preference segments. Average ratings were found to be different at well beyond the .001 level of significance, suggesting that the segments were indeed different in their preferences for this ethnic food product.

#### Results

Table 3 shows both direct and derived ideal scale values for each of the preference segments. Table 4 shows the distribution of difference scores between the two types of ideal points.

TABLE 3  
DIRECT & DERIVED IDEAL POINTS FOR TWO PREFERENCE SEGMENTS

Product Characteristic	Segment A			Segment B		
	Direct	Derived	Difference	Direct	Derived	Difference
A	3.6	4.0	+4	3.8	6.0	+2.2
B	3.6	6.0	+2.4	3.4	6.0	+2.6
C	4.3	6.0	+1.7	3.8	3.0	-.8
D	4.7	5.0	+.3	4.1	3.0	-.9
E	5.1	6.0	+.9	5.2	6.0	+.8
F	3.6	4.0	+.4	4.2	6.0	+1.8
G	4.2	4.5	+.3	4.4	6.0	+1.6
H	2.8	4.0	+1.2	2.6	1.0	-1.6
I	4.8	6.0	+1.2	5.0	5.5	+.5
J	4.9	6.0	+1.1	4.8	6.0	+1.2
K	4.0	3.5	-.5	4.6	6.0	+1.4

TABLE 4

#### DISTRIBUTIONS OF DIFFERENCE SCORES BETWEEN DIRECT AND DERIVED IDEAL SCALE VALUES

Difference	Segment A	Segment B	Total
+2.1 or more	1	2	3
+1.1 to +2.0	4	4	8
+0.1 to +1.0	5	2	7
0	--	--	--
-0.1 to -1.0	1	2	3
-1.1 to -2.0		1	1
-2.1 or more			

Several points of interest emerge from these tables:

1. Twelve of the twenty-two comparisons of direct vs. derived ideal points differ by more than one interval on the six-point scale; three differ by more than two intervals.
2. Respondents consistently understated their ideal points on the direct basis, as compared to derived. In only four of twenty-two cases did respondents indicate an ideal point that was higher than the one from cross-tabulations; the remainder were lower.
3. The amount of this understatement of ideal points is itself understated, based on the comparisons shown in Table 3. This table shows a "6" for derived ideal points whenever the dependent variable rating (i.e., overall evaluation) increased monotonically for all intervals of the independent variable. This means, though, that the "topping out" point for such a characteristic might well be at some value one, two, or even more intervals beyond the six intervals on the rating scale for that characteristic.
4. The only consistent pattern between the two preferences segments was in the case of the three characteristics which related to food color; both groups showed about 2.5 scale units difference of direct vs. derived ideal points for Characteristic B, about 1.2 scale units difference for Characteristic J, and about 0.8 scale units for Characteristic E. Other than for color, there seemed to be no really consistent pattern of differences between the two segments. It is not clear why color was so much more consistent than the other features; perhaps there is more agreement for visual cues than for gustatory (taste).

In summary, it seems reasonable to conclude that there was not really good agreement between direct and derived ideal points. The former tended to be understated in relation to the latter, raising the question as to which of the two is the more accurate. Unfortunately, there is no answer to this within the data available for this study; the best answer would require experimentation using several products which varied systematically in the levels of all attributes rated. Such an approach was well beyond the resources available for this study.

At the very least, this study suggests that future research should not rely too heavily on subjective ideal point ratings from respondents. These should be supplemented by cross-tabulations of product characteristics vs. overall evaluation ratings, to provide at least another estimate of the optimum level of each characteristic. The latter approach has the advantages of objectivity plus reference points on each attribute based on average ratings for each of the products on each attribute. Optimum levels for each attribute can then be determined by interpolation or extrapolation from measurable levels on the products being tested.

## References

- Seymour Banks, Experimentation in Marketing (McGraw Hill, 1965).
- F. M. Bass, E. A. Pessemier, and D. R. Lehmann, "An Experimental Study of Relationships Between Attitudes, Brand Preference and Choice," Institute Paper No. 307, Lafayette, Ind., Krannert Graduate School of Industrial Administration, Purdue University, 1971.
- J. D. Carroll and J. J. Chang, "Relating Preference Data to Multidimensional Scaling Solutions via a Generalization of Coombs' Unfolding Model," Bell Telephone Laboratories, Murray Hill, New Jersey.
- J. O. Eastlack, "Consumer Flavor Preference Factors in Food Product Design," Journal of Marketing Research, 1 (February, 1964), 38-42.
- Paul E. Green, "MDS & Related Technologies: Where Do We Go From Here?" Journal of Marketing, 39 (January, 1975), 24-31.
- Paul E. Green and F. J. Carmone, Multidimensional Scaling and Related Techniques in Marketing Analysis (Allyn and Bacon, Boston, 1970).
- D. Klahr, "A Study of Consumers' Cognitive Structure for Cigarette Brands," Journal of Business, 43 (April, 1970), 190-204.
- Richard J. Johnson, "Market Segmentation: A Strategic Management Tool," Journal of Marketing Research, 8 (February, 1971), 13-19.
- L. A. Neidell, "Physician Perception and Evaluation of Selected Ethical Drugs: An Application of Non-metric Multidimensional Scaling," dissertation, University of Pennsylvania, 1969.
- Allan D. Shocker and V. Srinivasan, "A Consumer-Based Methodology for the Identification of New Product Ideas," Management Science, 20 (February, 1974), 921-937.
- William L. Wilkie and Edgar A. Pessemier, "Issues in Marketing's Use of Multi-Attribute Attitude Models," Journal of Marketing Research, 10 (November, 1973), 428-441.
- Fred S. Zufryden, "A Zero-One Integer Programming Approach to Market Segmentation & Product Positioning," GSBA Working Paper, School of Business Administration, University of Southern California, 1976.

BEHAVIORAL MEASUREMENT OF THE RELATIVE IMPORTANCE OF ATTRIBUTE-RELATED INFORMATION CUES:  
THE CASE OF COLD BREAKFAST CEREALS<sup>1</sup>

John A. Quelch, University of Western Ontario

Abstract

An information display was used to investigate the relative importance of a set of attribute-related information cues in influencing trial and repeat brand choice behavior for a frequently purchased product category, adult-oriented (i.e., non-presweetened) cold breakfast cereals. Information on subjective attribute dimensions (including physical samples of the brands on offer) was included in the display as well as information on objective attribute dimensions.

Introduction

Myers and Alpert (1968 and 1977) have concluded that an attribute dimension may be important to a consumer but may not influence the brand decision-making process in a determinant sense if all brand alternatives are perceived by the consumer to be similar on that dimension. The distinction between importance and determinance emerged from efforts to improve the predictive power of the multiattribute model. For example, Wilkie and Weinreich (1972) reported that predictive power diminished beyond the inclusion of the two or three attributes with the highest determinance scores. However, while attribute importance weights and determinance scores may vary, the attitudinal nature of the data inputs to the model prevent the researcher from distinguishing those attributes sufficiently important or determinant to be used in decision-making from those which are not.

The pilot study reported here obtained behavioral measures of relative attribute importance through the use of an information display matrix of brand-attribute information. Characteristically, participants in information display studies elicit as few or as many items of information as they wish prior to selecting a brand. The researcher can, therefore, distinguish those attribute dimensions which are considered sufficiently important to be consulted before a brand selection is made from those which are not. Prior evidence (Jacoby, Chestnut, Weigl and Fisher, 1976) indicates that few consumers acquire all the information available to them before making their brand selections. In this study, determinant attributes are treated first as those used by a consumer in the decision-making process and secondly as those which result in a reduction in the number of brand alternatives under consideration in the process.

Method

Each subject was required to select a brand of adult-oriented cold breakfast cereal from among an array of six brands on offer in an information display. To assist in their decision-making, subjects could elicit information about the attribute profile of each brand.

The Sample

A non-random sample of 250 female homemakers was recruited at a shopping mall in Massachusetts. Potential

subjects were screened for recent purchase and consumption of cold breakfast cereal. In addition, they had to designate an adult-oriented brand when asked to name a favorite. Qualifying subjects were invited to make appointments to participate in a study investigating a new market research technique. Each consumer was told that she would receive ten dollars in cash and coupons for her cooperation. Of 341 prospective subjects, fourteen failed to qualify on the screening criteria. The remaining non-participants either declined to participate or failed to keep their appointments.

To avoid bias against inclusion of working women in the sample population, screening and task administration were conducted during evening as well as day-time hours and during one Saturday as well as three weekdays.

Subsequent analysis of the sample indicated that lesser educated, lower income, and elderly consumers were under-represented in terms of both local and national census figures. Aggregate results cannot, therefore, be considered nationally or locally representative.

Display Structure and Task Administration

The brands on offer were listed horizontally at the same eye level. However, because people characteristically read from left to right, it was thought that a bias towards selection of information about brands on the left side of the information display might occur. Since randomization of the order in which brands were listed for each subject would have created logistical difficulties, the potential problem was investigated through the use of two information displays in different rooms at the mall location. One display presented the brands in alphabetical order, the other presented them in reverse alphabetical order. Subsequent analysis indicated no significant differences in the results derived from each information display in terms of the mean numbers of information units elicited for each brand. This test may, however, have been inadequate for the third and fourth brands in the sequence, whose positions were simply reversed between information displays, particularly if both primacy and recency effects were operative.

Since people commonly read from top to bottom, the listing of attribute dimensions on the vertical dimension presented a similar problem of order effects. In this study, attribute dimensions were not listed on the information displays. Instead, subjects were presented on a rotation basis with cards which listed the available attribute dimensions in different sequences.

Six project personnel, who had been briefed as a group and had been involved in a feasibility test of the design, administered the information display task, recording the information elicited by each subject as she proceeded. Insofar as work schedules and minimization of subject waiting time permitted, personnel were rotated between the rooms containing the two information displays to limit the impact of any interviewer-related biases.

<sup>1</sup>The author acknowledges the support of the Marketing Science Institute in funding this study.

## Specification of Treatment Groups and Attributes

The specification of attribute dimensions is no easier a task in information display studies than in investigations of relative attribute importance using other methodologies. Subjects were assigned to one of two treatment groups which varied in the number of attribute dimensions available for elicitation. Subjects in the first treatment group (124 subjects) were presented with four information cues - ingredient information, nutrition information, physical appearance, and price. Subjects in the second treatment group (126 subjects) were presented with an additional information cue - the opportunity to elicit physical samples of the brands on offer in the exercise.

The nutrition information and ingredient information cues replicated the information presented on the boxes of the cereal brands on offer in the exercise. The format and specificity of information on these dimensions did not vary across brands. Price information represented an average of store prices in the area from which the sample was drawn, standardized for a twelve ounce package. This package size was specified on each of the price information cards.

The physical appearance dimension included a verbal description of the size, color, and texture of the cereal pieces, and a life-size drawing which enabled subjects to tell whether a particular brand was in the form of flakes, nuggets, or puffs. Photographs were not used to avoid brand identities being revealed. It is possible that the number of subjects who elicited the physical appearance dimension may have been unrealistically reduced as a result.

Physical samples of the brands on offer were available in bowls, covers on which were removed if subjects elicited the "physical sample to look at and/or to taste" attribute dimension. The number of subjects in the second treatment group who elicited physical sample information may have been unrealistically limited. Some subjects may have considered the sampling procedure unhygienic, others may have been disinclined to sample the brands in a semi-public environment. Still others may simply not have been hungry. In addition, subjects were not able to sample the brands with milk or sugar.

With the exception of price, these information cues each provided subjects with information about several attributes. For example, the nutrition information dimension provided subjects with the calorie content, protein content, and vitamin and mineral content of each brand. The rationale for not decomposing these dimensions into their constituent attributes stemmed from one of the main study objectives. This was to investigate possible variations in the relative importance of attribute dimensions among each other, conditional upon whether the purchase occasion represents a trial or a repeat purchase. In a repeat purchase situation, the consumer has had an opportunity to experience the product or brand in question. In the case of cold breakfast cereals, it was hypothesized that the importance of physical sample information to the repeat buyer would be such that the importance of the other attribute dimensions on which performance information was available to both trial and repeat buyers would be significantly reduced in the repeat purchase situation. The information available to the two treatment groups was therefore intended to represent those attribute dimensions on which performance information is available to the consumer making a trial and a repeat purchase respectively. The attribute dimensions included in each set were not evoked from subjects in this study but were based upon the results of proprietary consumer research conducted by a major cereal manufacturer. It should be noted that the simulation

did not consider the impact of differential information recall, personal influences, and other distortions common to naturally occurring purchase situations.

Recognizing that actual physical samples of brands on offer in the display would constitute a multiattribute information cue, it was decided that the attribute dimensions common to both treatment groups should not be decomposed into their constituent attributes. To have done so would, furthermore, have resulted in a preponderance of nutrition or ingredient related attributes, potentially affording them an unnatural prominence and enhancing the possibility of consumer learning biasing the information elicited. In addition, Arch, Bettman, and Kakkar (1978) have indicated that in real purchase situations a consumer may acquire information on several attributes simultaneously by examining the nutrition label. Decomposition of the nutrition information (and other) dimensions might have had the effect of increasing acquisition time and thereby reducing the total quantity of information acquired. The biases which presentation formats impose upon the external validity of the information acquired in display studies merit further investigation. The search time for each information unit is equalized. The absolute and relative length of time required to search information on different attributes in the real world may not be accurately replicated. Since information search time may influence the quantity of information and the nature of attribute information acquired, the accuracy of the results of information display studies may justifiably be questioned on this basis. In particular, the display of information available may actually stimulate demand for information which, in a real purchase situation, might not have influenced choice behavior.

In contrast to the Jacoby, Chestnut, Weigl, and Fisher (1976) study, brand name was not included as an information cue. If brand name had been available and had been the only cue elicited by a subject, it would not have been possible to determine whether non-selection of additional attribute dimensions was due to their relative unimportance or to the fact that the information which they could have provided was subsumed in the brand name dimension. Similarly, selection of additional attribute dimensions might have been due to the complexity of the information associated with them, or to the difficulty of storing this information in long-term memory. Additional attribute dimensions consulted beyond brand name could, therefore, have been not the most important but those on which consumers found it relatively more difficult to store information. In addition, if brand name had been available as an information dimension, some consumers might have sought more information on those brands with which they were less familiar, and vice versa. For these reasons and because brand name was not included as an information dimension in this study, the information elicited by subjects may not represent the information which consumers with purchase experience in the cereal category would actually acquire when making a brand purchase in a supermarket.

## Specification of Brands

Six brands of cold breakfast cereal were on offer in the brand selection exercise. The brands were real, but were identified by code letters (L through Q) rather than brand name. To discourage subjects from selecting attribute information in order to identify their favorite brands, they were told that the brands on offer were not currently available in New England. Clearly, the attribute dimensions most useful in brand identification may not necessarily be those most important to the consumer.



One constraint upon the choice of brands was the need to avoid identifiable brands, since physical samples were available as an information cue to subjects in the second treatment group. Otherwise, subjects might have believed that they had identified particular brands and might have made their brand selections on the basis of brand name impression rather than the attribute information available from the information display.

After consultation with marketing research executives of a major cereal manufacturer, six brands were selected to represent the adult-oriented segment of the cold breakfast cereal market. Representativeness of brands in terms of product form, price, nutritional and ingredient content, was essential if measures of the relative importance of these attribute dimensions were to be generalizable rather than specific to the brand set. Industry studies classified two of the six brands (Kellogg's Corn Flakes and Rice Chex) as all-family, two (Total and Special K) as fortified, and there was, in addition, one bran (Post Raisin Bran), and one granola (Quaker Natural). In terms of volume, the six brands represented about twenty percent of the cold breakfast cereal market.

#### Procedural Constraints

Ordinarily, in information display studies, subjects have elicited information on a cell-by-cell basis. Since this study focused on investigating the relative importance of attribute information, subjects were constrained to acquiring information attribute-by-attribute across brands, and they were not permitted to return to an attribute dimension a second time to acquire additional information.

This procedure is compatible with a lexicographic model of information acquisition which assumes that consumers examine several brands on the most important attribute and, in the event of a tie, examine the remaining brands on the second most important attribute and so on until a preferred brand emerges. The lexicographic model suggests that the relative importance of attributes be delineated in advance of the selection process (Wright, 1972) and may, therefore, be appropriate for the study of frequently purchased product categories (such as breakfast cereals) in which most subjects have prior purchase experience.

Information display studies have indicated that more consumers use an attribute-by-attribute approach to information acquisition than a brand-by-brand approach (Bettman and Kakkar, 1977; Jacoby, Chestnut, Weigl, and Fisher, 1976; Reilly, Holman, and Evered, 1976). In an eye movement study, Russo and Rosen (1975) also found a greater propensity to information acquisition attribute-by-attribute. While the literature supports the attribute approach, Bettman and Jacoby (1976) have pointed out that, because product information is organized by brand on the supermarket shelf, some consumers who prefer to use an attribute-by-attribute approach may, in fact, use a brand-by-brand approach to information acquisition. The common procedure used in this study may have constrained some consumers to follow an information acquisition strategy with which they were unfamiliar or uncomfortable. In these cases, information processing time may have been unrealistically increased or the quantity of information acquired may have been unrealistically curtailed.

Subjects had to consult the information cards for all six brands on the first attribute dimension which they selected. The purpose of this constraint was to encourage subjects from the outset to look at the information display as a whole rather than to concentrate on the left-hand side, for example. In addition, one research objective was to ascertain the percentage

of consumers whose favorite brand of cereal was one of the six brands on offer in the exercise who, in fact, chose the favorite brand<sup>2</sup>. To do this, it was necessary to ensure that all subjects were exposed to at least one item of information about all six brands.

This constraint may have affected the brand selection process of those consumers who had in mind to acquire a fixed quantity of information or to spend a fixed period of time on the task prior to making a brand selection. Some subjects might have consulted fewer brands on the first attribute dimension. Alternatively, the first attribute dimension selected might not have been the most important, but, rather, one on which the performance of six brands was perceived as being reviewable in a shorter period of time.

Beyond the first attribute dimension, the constraint did not apply. Subjects could choose the brand or brands for which they elicited information on each additional attribute dimension. They were not, therefore, required to take potentially redundant attribute information on brands which they might have eliminated from further consideration on the basis of information already acquired.

Once a subject selected an attribute dimension, she was asked to remove the cards for those of the six brands on which she wished to see the information before reviewing any of these cards. While this constraint may have increased information acquisition time, its purpose was twofold. First, it counteracted the advantage which brands on the left side of the information display might have had if a subject, proceeding from left to right across a row of attribute information, had been permitted to review each card in turn. Secondly, the validity of investigations into whether selection of particular attribute dimensions resulted in differential reduction in the number of brands considered on the next dimension elicited would have been impaired if subjects had been able to review information cards one by one - the reason being that the number of cards actually removed might have been fewer than the number of cards which the subject originally intended to take based on brand information elicited on the previous attribute dimension.

Subjects in this study were not required to turn information cards face down once they had looked at them. In a study focusing on relative attribute importance, it would have been difficult to determine whether back referencing was an indication of the importance of the attribute dimension, consumer forgetfulness, or the relative difficulty of retaining the information. Accordingly, subjects in this study were permitted to freely refer back to information already acquired.

#### Compensation

To introduce a sense of motivational realism to the task similar to that present in naturally occurring purchase situations, and to discourage elicitation of information on the basis of perceived social acceptability, subjects were told that they would receive two fifteen cents-off coupons towards future purchases of the brand which they selected.

<sup>2</sup>These results are reported in John A. Quelch, "Explaining and Predicting Brand Choice Behavior: An Application of Information Display Research", Working Paper No. 181, School of Business Administration, The University of Western Ontario. In fact, of ninety-six subjects who had the opportunity to select their favorite brands, thirty-four did so, a proportion significantly (at the 0.05 level) higher than the one-sixth or sixteen subjects which would have been expected by chance.

It was believed that the compensation level could influence the perceived importance of the task and, therefore, the quantity and nature of information elicited. A free box of each subject's selected brand was rejected because it removed any incentive to elicit price information, other than perhaps to discover the highest-priced brand. Measurements of the relative importance of the attribute dimensions included in the exercise would have been rendered unrealistic.

It was believed that a fifteen cents-off coupon was within the normal range of coupon values for cold breakfast cereals. No more than two coupons could be offered since brand choice decisions are, in reality, discrete, and consumers purchase on average between one and two boxes of cold breakfast cereals at any one time. To have offered more than two coupons might have caused subjects to become unrealistically risk averse in their information acquisition, attribute elicitation, and brand selection strategies.

### Results

In this section, the quantity of information acquired by subjects in terms of number of attribute dimensions consulted and number of information units elicited is first reported together with an examination of the relationship between quantity of information available and quantity of information acquired across treatment groups. Secondly, three measures of relative attribute importance are applied to the data.

#### Acquisition of Information

No subject in either treatment group elicited all the information units available. Indeed, 23 percent of subjects were willing to base their brand selection decisions on only six information units, the minimum number which they were permitted to take.

Across the whole sample, each subject consulted an average of 2.26 attribute dimensions. A larger percentage (40 percent) of subjects selected information on two attribute dimensions than on any other number of attributes. The results indicated that the majority of consumers were content to use only one or two attribute dimensions as the basis for making brand selections.

Comparing the results for subjects in both treatment groups, the quantity of information acquired was found to be unrelated to the quantity of information available. The mean numbers of information units acquired were respectively, 9.79 for subjects in the first treatment group (with 24 units available), and 9.34 for subjects in the second group (with 30 units available). The corresponding mean numbers of attributes consulted were 2.25 and 2.27. In neither case was the difference between the two means statistically significant.

It is possible that the incremental six units of information were insufficient to make meaningful any comparison between the two treatment groups in terms of the quantity of information available, particularly since the additional attribute dimension available to subjects in the second treatment group (physical sample information) was, in fact, an expanded version of the physical appearance dimension and may have been regarded by subjects as substitute rather than independent information.

The numbers of information units acquired by subjects in this study were similar to those reported by Jacoby, Chestnut, Weigl, and Fisher (1976) for an information display study in which subjects were also required to select a brand of cold breakfast cereal. The two studies differed in the number and nature of the brands and attributes included, in the nature of the task, and

in the level of compensation offered to participants. Interestingly, however, half of the Jacoby et al. sample acquired eleven or fewer information units, compared to 74 percent of subjects in this study. The tendency of subjects in this study to acquire fewer information units may be explained by the fact that the information dimensions were not as decomposed as in the Jacoby et al. study. For example, nutrition information was presented as a single attribute dimension in this study, whereas it was decomposed by Jacoby et al. into the constituent nutrients. An information unit in this study tended, therefore, to represent a larger "chunk" of information than an information unit in the Jacoby et al. study.

#### Behavioral Measures of Relative Importance

Three different, though not independent, measures of the relative importance of information on attribute dimensions were derived from the data collected in the brand selection exercise:

Selection or non-selection of an attribute dimension. This distinction enables those attribute dimensions to be identified which were perceived by each subject to be sufficiently determinant to warrant investigation. Other attribute dimensions, not consulted, may also have been perceived to be determinant (i.e., both important and differentiated across brands) but not to the level necessary to justify their inclusion in the decision-making process. Perceived determinance is a necessary but not a sufficient condition for an attribute to be consulted. The aggregate frequency with which each of the five information dimensions was elicited by members of the two treatment groups is indicated in the following tabulation:

	Treatment Groups	
	No Physical Sample (NPS) n=124	Physical Sample (PS) n=126
Ingredient Information	91	66
Nutrition Information	91	70
Physical Appearance	50	20
Price	47	44
Physical Sample	N/A	86
Total Elicitation Occasions	279	286

The results, assuming sample homogeneity, indicate that nutrition information and ingredient information were most frequently perceived as determinant attribute dimensions in the trial purchase situation, while physical sample information was most often perceived as determinant in the repeat purchase situation.

Given that the total number of elicitation occasions did not increase significantly between treatment groups, consultation of the physical sample dimension clearly occurred at the expense of the other four. The results enable the comparative changes in importance of these attribute dimensions to be assessed. For example, the share of total elicitation occasions held by the physical appearance dimension declined substantially between treatment groups, while the share held by the price dimension remained constant. The differential changes in elicitation occasion shares probably reflect, in part, the relative degree to which information on each of the four common attribute dimensions was perceived to be obtainable from information on the physical sample dimension. For example, physical samples of breakfast cereals may provide some consumers with information about their principal ingredients but do not apparently provide sufficient information on the price dimension to cause a decrease in the frequency with which this attribute is consulted.

Order of Attribute Dimension Selection. Table 1 shows the frequency with which each information dimension was elicited by subjects in the two treatment groups for each of five possible selection sequence positions. A similar measure of relative attribute importance has been used in an information display study reported by Holbrook and Maier (1978). Conclusions drawn from the frequency of elicitation data can be amplified if the rank order data are examined. For example, while the overall frequency of elicitation of ingredient information fell by 27 percent between the two treatment groups, the frequency with which this attribute dimension was elicited first fell by 42 percent. In contrast, the overall frequency of elicitation of nutrition information fell by 23 percent, but the frequency with which it was elicited first fell by only 18 percent. Nutrition information appeared, therefore, to be less susceptible to displacement than ingredient information, given the availability of the physical samples.

The credence attached to this measure of relative importance rests upon the assumption that consumers will first elicit information on the attribute which is perceived a priori to be the most important with a view to minimizing the duration of the decision-making process. In this regard, three problem areas should be noted. First, for the consumer whose decision-making process involves a distinction between narrowing the range of brand options to a set of acceptable alternatives, and then choosing from among the brands in this set, the first attribute dimension selected may not be the most important in contributing to the final decision - particularly in the case of a consumer who uses different attribute dimensions at different stages of the decision-making process. Secondly, the order in which attribute information is acquired throws no light on the intended use of the information - in particular, whether the subject wishes to treat the information in a threshold or maximizing sense. Third, in the absence of prior purchase experience and prior definition of purchase goals, subjects may choose attribute dimensions which they believe to be important but which prove in fact to be less important in facilitating a brand choice decision than other attribute dimensions elicited later in the rank order. In brief, a priori and post hoc assessments of relative importance may be divergent.

To investigate this third problem area, each subject in the study was asked upon completion of the exercise to allocate one hundred points among those attribute dimensions which she consulted to reflect the relative importance of each in facilitating the brand selection process. For each subject, a Spearman nonparametric rank order correlation coefficient was computed between the rank order in which the attribute dimensions were selected, and the rank order in which points were allocated to each of the same dimensions. The mean coefficient across all subjects was 0.71. The reliability of the points allocation measure may have been distorted by socially acceptable response patterns (for example, allocating a disproportionate percentage of

points to nutrition information), and by cognitive consistency pressures (prompting subjects to reflect the rank order of elicitation in their points allocations). With these reservations, the correlation statistic does suggest that, among consumers accustomed to making brand choice decisions within a particular product category, the rank order of elicitation may be an appropriate surrogate for relative importance.

Determinance Measure. In addition to distinguishing those attribute dimensions sufficiently determinant to warrant consultation from those which were not consulted, a method of identifying the relative determinance of those attribute dimensions which were consulted was developed.

The determinance of an attribute dimension was operationally defined as the degree to which its elicitation caused a reduction in the number of brands under consideration (measured in terms of the number of brands on which the next attribute dimension was elicited). A feasibility test had indicated that most subjects could be expected to "pyramid" towards their final brand selections in this manner. Since subjects were constrained to receive information on all six brands for the first attribute dimension, a mean determinance score could be calculated for each attribute dimension by treatment group on the basis of the average number of brands on which the second attribute dimension was elicited. Thus, the lower the mean score, the more deterministic the attribute dimension in question. In those cases where the first attribute dimension selected was the only one selected prior to a brand choice being made, a determinance score of zero was recorded. The following tabulation presents mean determinance scores by information dimension and by treatment group:

	Mean Determinance Scores	
	NPS	PS
Ingredient Information	2.125	2.438
Nutrition Information	2.333	2.357
Physical Appearance	3.600	2.900
Price	3.714	2.667
Physical Sample	N/A	1.472
Aggregate	2.419	2.183

The results indicate that for subjects in the second treatment group (representing the repeat purchase condition), physical sample was the most determinant attribute dimension, in part reflecting the fact that 42 percent of subjects who selected this attribute dimension first did not seek any further information. Subjects who selected this attribute dimension first selected information on the second attribute dimension for an average of only 1.472 brands, compared to 2.357 brands when nutrition information was selected first (t statistic for difference between means significant at 0.05 level).

Two reservations regarding this determinance measure should be noted. First, the degree to which consultations of attribute dimensions resulted in reductions in

TABLE 1  
INFORMATION CUE ELICITATION OCCASIONS BY SELECTION SEQUENCE POSITION AND BY TREATMENT GROUP

	0		1		2		3		4		5	
	NPS	PS	NPS	PS	NPS	PS	NPS	PS	NPS	PS	NPS	PS
Ingredient Information	33	60	56	32	26	20	7	8	2	6	0	0
Nutrition Information	33	56	51	42	32	22	6	5	2	1	0	0
Physical Appearance	74	106	10	10	21	4	17	4	2	1	0	1
Price	77	82	7	6	17	17	13	16	10	4	0	1
Physical Sample	N/A	40	N/A	36	N/A	33	N/A	15	N/A	2	N/A	0

NOTE: To be read (for example): "33 subjects in the first treatment group did not elicit ingredient information, 56 elicited ingredient information first, etc."

the number of brands under consideration when the attribute dimensions were consulted other than first in order of elicitation were not considered. Differences in the nature of information previously consulted would have made it impossible to compare subjects on a common basis. Nevertheless, the determinance measure was deficient in that it was based on only a portion of the data available. Secondly, the determinance measure assumes that the pattern of processing (in particular, the number of brands under consideration) is solely a function of previously acquired attribute information. Predisposition regarding search time and number of information units to be acquired, given the nature of the decision, probably vary across individuals. Such individual differences have been considered randomly distributed throughout the sample.

#### Summary and Implications

Although information display exercises are behavioral in nature, artificialities clearly exist in the manner in which the information is presented and in the procedural constraints which are imposed upon subjects in order to facilitate subsequent analysis. The impact of these artificialities on external validity requires further investigation. In addition, the information display approach offers no advance over other techniques in the area of specification of the attribute dimensions to be included.

However, the use of information displays to measure relative attribute importance offers two principal advantages over traditional attitudinal measures. First, the behavioral approach permits motivational realism to be incorporated into the data collection process, so that the propensity of subjects to provide what they perceive to be socially acceptable response patterns may be reduced.

Secondly, by permitting each subject to determine how many information units are to be acquired prior to a brand selection decision, information display exercises provide data which enable the researcher to distinguish those attribute dimensions perceived by subjects to be sufficiently important to warrant investigation from those not so perceived. The distinction cannot be drawn among a set of attribute dimensions which are rank ordered directly or according to attitudinal importance weights. In addition, information display exercises permit differences in the relative importance of attribute dimensions under varying conditions of information availability to be identified.

The relevance to policy makers of whether attribute information available to consumers is acquired has been emphasized by Jacoby, Chestnut, and Silberman (1977), and by Scammon (1977). The question is also of significance to marketers who must ensure that their products are superior on those attribute dimensions which target consumers consider sufficiently important to warrant inter-brand comparisons in their decision-making processes.

#### References

David C. Arch, James R. Bettman, and Pradeep Kakkar (1978), "Subjects' Information Processing in Information Display Board Studies", in H. Keith Hunt, ed., Advances in Consumer Research, Vol. 5, Association for Consumer Research, 555-560.

James Bettman, and Jacob Jacoby (1976), "Patterns of Processing in Consumer Information Acquisition", in Beverlee B. Anderson, ed., Advances in Consumer Research, Vol. 3, Association for Consumer Research, 315-320.

James Bettman, and Pradeep Kakkar (1977), "Effects of Information Presentation Format on Consumer Information Acquisition Strategies", Journal of Consumer Research, 3, 233-240.

Morris B. Holbrook, and Karl A. Maier (1978), "A Study of the Interface Between Attitude Structure and Information Acquisition Using a Questionnaire - Based Information Display Sheet", in H. Keith Hunt, ed., Advances in Consumer Research, Vol. 5, Association for Consumer Research, 93-98.

Jacob Jacoby, R.W. Chestnut, and William Silberman (1977), "Consumer Use and Comprehension of Nutrition Information", Journal of Consumer Research, 4, 119-128.

-----, -----, Karl C. Weigl, and William A. Fisher (1976), "Pre-Purchase Information Acquisition: Description of a Process Methodology, Research Paradigm, and Pilot Investigation", in Beverlee B. Anderson, ed., Advances in Consumer Research, Vol. 3, Association for Consumer Research, 306-314.

James H. Myers, and Mark I. Alpert (1968), "Determinant Buying Attitudes: Meaning and Measurement", Journal of Marketing, 32, 13-20.

-----, ----- (1977), "Semantic Confusion in Attitude Research: Saliency vs. Importance vs. Determinance", in William D. Perrault Jr., ed., Advances in Consumer Research, Vol. 4, Association for Consumer Research.

M. Reilly, R. Holman and R. Evered (1976), "Individual Differences in Information Processing: An Exploratory Report", in Proceedings, 8th Annual Conference, American Institute of Decision Sciences.

Edward Russo, and Larry D. Rosen (1975), "An Eye Fixation Analysis of Multi-Alternative Choice", Memory and Cognition, 3, 267-276.

Debra L. Scammon (1977), "Information Load and Consumers", Journal of Consumer Research, 4, 148-155.

William L. Wilkie, and Rolf P. Weinreich (1972), "Effects of the Number and Types of Attributes Included in an Attitude Model: More is Not Better", in M. Venkatesan, ed., Proceedings, Third Annual Conference, Association for Consumer Research, 325-340.

Peter L. Wright (1972), "Consumer Judgement Strategies - Beyond the Compensatory Assumption", in M. Venkatesan, ed., Proceedings, Third Annual Conference, Association for Consumer Research, 316-324.

THE STABILITY OF RESPONSES OBTAINED BY FREE ELICITATION:  
IMPLICATIONS FOR MEASURING ATTRIBUTE SALIENCE AND MEMORY STRUCTURE<sup>1</sup>

Jerry C. Olson, Pennsylvania State University  
Aydin Muderrisoglu (student), Pennsylvania State University

Abstract

This paper examines the stability over time of consumers' verbal responses in a variety of free elicitation tasks and the test-retest reliability of several measures derived from those responses. The paper concludes with a discussion of the potential usefulness of the free elicitation procedure for identifying salient product attributes and for indicating the content and organization of product knowledge structures as stored in memory.

Introduction

A variety of marketing research issues involve determining the content of consumers' cognitive structures for products, brands, or other concepts. A common example is the need to identify the product attributes consumers believe to be salient (Sampson and Harris, 1970), important (Cohen, Fishbein, and Ahtola, 1972), or determinant (Myers and Alpert, 1968; 1977). Once identified, these product characteristics are typically incorporated into a structured questionnaire. Consumer responses are then obtained in a mass administration and are usually combined via some multi-attribute model to predict consumer attitudes or choice behavior (cf. Day, 1972; Wilkie and Pessemier, 1973). A second example involves a related, but more basic issue--namely, how consumers encode, organize, and store in memory their knowledge about a product and its characteristics (Johnson and Russo, 1978; Olson, in press). Determining the content and organization of a consumer's product knowledge structure requires procedures and techniques by which a researcher can identify (a) the cognitive form or symbol the consumer uses to represent a product characteristic, (b) the meaning (comprehension) "attached to" that symbol, (c) the relative "importance" (vis-a-vis some criterion) of that knowledge representation, and (d) the interrelationships or structure among various product knowledge characteristics. This paper describes one such procedure--free elicitation--and explores the stability of the data generated by it.

A number of approaches and techniques have been used in attempts to identify the salient or "important" dimensions of cognitive structure. Perhaps most commonly reported in the marketing literature is the use of a loosely-structured individual or group interview, the details of which unfortunately, are seldom described. More structured and therefore more replicable procedures exist, but these have been used much less frequently in marketing research. Examples include the Kelly (1955) repertory grid (Frost and Baine, 1967; Sampson, 1972) or modifications thereof (Wilson and Dover, 1975), cued or "controlled" elicitations (Fishbein, 1967), and free recall (Green, Wind and Jain, 1973; Johnson and Russo, 1978). An exhaustive review of the pros and cons of these techniques is beyond the scope of this paper, as is a detailed rationale for the selection of the elicitation procedure advocated here. Suffice it to say that recent developments (Abelson, 1973, 1976; Calder, 1978) indicate that peoples' cognitive

structures contain a variety of types of concepts including abstract, attribute-specific dimensions as well as more concrete, visual concepts such as images (Mower, 1977) and scripts (Shank and Abelson, 1977). Thus, it was deemed important that the procedure to be investigated was sufficiently "open" and nondirective that the full variety of cognitive elements maintained in a memory structure could be evidenced. This paper examines a free elicitation procedure in which respondents are free to say anything and everything that comes to mind when presented with a stimulus probe cue.

In cognitive research, a common method for investigating the content and organization of memory structure is to ask the respondent to freely recall previously learned information (cf. Buschke, 1977). The free elicitation procedure of interest here is essentially the same, but for two important differences. First, in free recall the probe cue that triggers responding is a phrase such as, "Tell me the words we learned yesterday..(or, last week)."  
In free elicitation, the probe tends to be more general and non-time specific--an attempt to "trigger" or activate a particular structure of stored knowledge--for example, "Tell me what comes to mind when I say Ford LDT!,"  
Second, in contrast to free recall, the particular learning event or the content acquired from a specific learning experience tends not to be of interest in free elicitation. Rather, the researcher is primarily interested in the content and organization of an existing structure of knowledge stored in semantic memory (Olson, in press).

Once generated, the responses elicited in either free recall or free elicitation tasks can then be examined as evidence of memory content and organization. A wide variety of measures and indices can be developed from free response data. For example, one could examine the ordering of the elicited concepts, the clustering patterns of concepts, the rate of elicitation, the time taken between responses, etc. Moreover, one can examine the nature of individual responses. For example, is the elicited concept an abstract product characteristic such as convenience, or is it more concrete (i.e., more visual or imaginal)? Although free response procedures have been used occasionally in cognitive research, there seems to be little published evidence regarding the stability of such data, or the reliability of measures derived from such data, or, indeed, the validity of the approach for indicating the content and organizational characteristics of knowledge structures. Thus, before the free elicitation procedure can be proposed as possibly useful for indicating consumers' memory content and structure, the psychometric qualities of the data need to be established, beginning with issues of stability and reliability.

A variety of approaches to establishing reliability have been identified (cf. Nunnally, 1967; Silk, 1977) and each has specific advantages and particularly appropriate applications (Peter, 1977). Underlying all theories of reliability, however, is the idea of "consistency" or "stability" of response, and this is the focus of interest here. The present study involves a variety of single measures or indices (of possibly different constructs), most of which are not yet fully explicated. Thus, the reliability approach appropriate for this somewhat primitive state of measurement must focus on the test-retest reliability of individual measures. The hope, of course, is that the free elicitation procedure is capable of producing reasonably stable sets of responses and that

<sup>1</sup>This research was funded by a Research Initiation Grant to the first author from Pennsylvania State University. The authors are grateful to Rajesh Kanwar and Sik Shum for their assistance in data collection.

important aspects of these responses can be described by reliable measures. If so, subsequent attention can be devoted to developing theoretically meaningful interpretations of the measured constructs in terms of memory, content and organization.

## Method

### Research Overview

Naturally occurring cognitive structures were the focus of this study. That is, there were no experimental manipulations of the product knowledge that consumers had stored in memory. To produce additional data per subject and to allow examination of differential effects of factors such as product familiarity and ego-involvement, free responses were elicited vis-a-vis three product categories--toothpaste, ball-point pens, and blue jeans.<sup>2</sup> Because the three products examined were common and frequently purchased by the subject population, subjects were presumed to have previously acquired rather stable knowledge structures for each of the product categories, derived from their fairly extensive past experience.

The knowledge structures associated with each product were examined by presenting subjects with three different types of stimulus cues (memory probes). The "broadest" probe cue was the product category name (e.g., "toothpaste"). A somewhat more specific probe involved presenting subjects with a combined product and purchase situation stimulus. The third type of probe was a specific brand name. Consumers were instructed to state "all the things that come to mind" each time the researcher mentioned a probe word or phrase. In order to examine the reliability of the measures of the concepts generated in these free elicitation tasks and the stability of the concepts themselves, consumers returned after one week and repeated the elicitations.

### Subjects

Thirty male and female undergraduate business students enrolled in the introductory marketing course volunteered to participate in the study in return for a \$4.00 cash payment.<sup>3</sup> Subjects were run individually, and each free response session required about 45 to 50 minutes with a subject.

### Elicitation Procedures

After a subject entered the study room, he/she was directed to a seat across the table from the researcher. On the table was a cassette tape recorder. After introducing himself and stating that the study concerned the subject's reactions to several common products and brands, the researcher handed the subject a written copy of the instructions for the first part of the study (see Appendix). To insure that each subject acquired a clear and thorough understanding of the free response procedures, the subject was asked to follow along on his copy while the researcher read the instructions aloud, verbally emphasizing important points (critical parts of the instructions are underlined in Appendix). In essence, the instructions asked the subject to verbally state, as

rapidly as possible,<sup>4</sup> any and all thoughts that came to his or her mind when the experimenter said the name of a product or brand. Subjects were told that it was too difficult to write down the responses and that, for accuracy, the responses would be recorded (but kept anonymous).

Then the recorder was turned on and the researcher said, "OK; tell me the things that come to mind when I say. . . softdrinks." This first response elicitation constituted a "warm-up" task for the subject to become familiar with the procedure and these responses were not analyzed. If the subject was too verbal or analytical in describing or justifying single thoughts (an occasional problem), the researcher pointed out that short phrases descriptive of each thought were sufficient. Then the researcher began the elicitations of interest by stating the name of one of the three products of particular concern--ballpoint pen, toothpaste or blue jeans. After the subject had stated his elicited thoughts for the first product, the researcher stated the names of the second and the third products and tape recorded the elicitations produced for each. The order of the three products was randomly determined for each subject. After each subject had freely responded to the three general product categories, the researcher stopped the recorder and told the subject to complete one of three two-page questionnaires (one for each product category, randomly selected). Each instrument contained questions about product usage, self-confidence, product importance, etc.

Subjects were then given the instructions for the second part of the study (see Appendix). Here, subjects were asked to verbally state the concepts and thoughts that came to mind when considering buying a product for their personal use, when told the product's name. The recorder was turned on and the researcher stated a product/personal use purchase situation cue for the same three products as before, one at a time in random order, and the subject's freely elicited responses were recorded for each. Then the recorder was stopped and the subject completed in random order the two-page questionnaires for the remaining two products.

Then, the subject was handed the instructions for the third part of the study (see Appendix) and was asked to follow them as they were read aloud by the researcher. The elicitation procedure was similar to that used previously except that the subject was to respond to specific brand names. Two well-known brand names were selected for each of the three products--Crest and Ultrabrite, Bic and Cross, Levi's and Wrangler.<sup>5</sup> Subjects were to verbally state the thoughts that come to mind when presented with each of these six brand names.

Following the elicitations, a sorting procedure was used to measure subject's ego-involvement with the three products and attitude and purchase intention measures were taken. First, each subject was given one of three decks of 5 x 7 index cards (chosen at random). On the cards in each deck were printed all of the locally-available brand names for that product. The subject was asked to sort through the card deck and remove those brands he/she had not heard of before. Then, the subject was told to examine the remaining brands/cards and sort them into three piles or categories, corresponding to the latitudes of

<sup>2</sup> These three products were selected essentially arbitrarily to meet the criteria of common purchase and use by a student population.

<sup>3</sup> Clearly, the external validity of this sample is limited. Rather, the purpose of this study is to demonstrate the potential stability of free elicitation data in the context of a particular subject population, not to demonstrate its generalizability to a wider population.

<sup>4</sup> The speeded aspect to the elicitation procedure is critical. The notion is that rapid responses are less subject to biases (e.g., social desirability and intentional distortion), and thus reflect more accurately the content of memory.

<sup>5</sup> These brand names were selected arbitrarily except for the criterion that they be familiar to most subjects.

acceptance, rejection, and neutrality (Sherif, Sherif and Nebergall, 1965).<sup>6</sup> Subjects then estimated the percentage of their product purchases over the past year that were devoted to each brand in the acceptance region. Finally, subjects indicated their attitude ( $A_o$ --5-point bad-good scale) and purchase intention ( $BI$ --7-point very unlikely-very likely scale) toward each familiar brand. After subjects finished the sorting and rating tasks with one product, the same procedures were repeated for the other two products, again in random order.

At the conclusion of the first session, subjects were given one dollar and told to return to the same room at the same time and day exactly one week later for the conclusion of the study. Subjects were advised of the importance of attending the second session and were reminded that they would receive the remaining \$3.00 of payment at that time. When each subject arrived for the second session, the study procedures were repeated exactly, with the same products and brands, except that the order of products and brands within each stage of the replication was randomly determined.

### Results and Discussions

The general purpose of this study was to determine the stability of freely elicited concepts and the reliability of summary measures of the concepts. In the context of this study, stability means that a probe cue (a word spoken by the researcher) presented to consumers at two points in time elicited similar responses. The reliability of measures of those responses is indicated by test-retest correlation coefficients. The results are discussed separately for the elicitations produced by each type of probe cue.

### Product Category Probes

Table 1 presents the means and reliabilities for selected measures of the concepts elicited in response to the product name cues. The total number of concepts elicited for the toothpaste and ballpoint pen probes was reasonably reliable (.77 and .74), but was less so (.47) for blue jeans. Time of response yielded only moderate reliabilities ( $r$ 's = .43 to .50), as did the rate of elicitation (.21 to .73). The degree to which similar concepts were elicited in the test-retest tasks was about the same for all three products (47% to 52%), and the rank order correspondence of these similar concepts over the two elicitations was positive but weak (.09 to .23). For all three product categories, abstract product attributes constituted approximately 70% of the similar elicited concepts, while the remaining 30% were more concrete.<sup>7</sup> The rank order correspondence among the abstract concepts was weakly positive (.20 to .29).

### Product/Purchase Context Probes

The stability/reliability measures for the concepts elicited in response to the combined product/purchase context probes are presented in Table 2. The reliabilities of the number of concepts elicited were reasonably strong (.54 to .77) and the reliabilities for time taken and rate of elicitation were moderate (.35 to .54 and .32 to .64). The degree to which the same concepts were elicited in the two trials was about the same for the three products (54% to 60%), and the rank order correspondence of the similar elicited concepts was positive and weak (.22 to .34). Again, abstract product attributes accounted for approximately 70% of the similar elicited concepts and the elicitation rank order correspondence between these product dimensions was moderate (.28 to .45).

TABLE 1  
MEANS AND RELIABILITIES FOR SELECTED MEASURES OF THE STABILITY OF FREE ELICITATIONS TO PRODUCT CATEGORY PROBES

Selected Measures	Product Category Probe Cue					
	Toothpaste		Ballpoint Pens		Blue Jeans	
	T <sub>1</sub>	T <sub>2</sub>	T <sub>1</sub>	T <sub>2</sub>	T <sub>1</sub>	T <sub>2</sub>
1. Number of Elicited Concepts	7.8	8.4	8.7	7.7	9.6	8.4
Reliability <sup>a</sup>	(.77)		(.74)		(.47)	
2. Time Required for Elicitation (sec.)	42.5	41.6	47.0	37.2	54.9	44.3
Reliability <sup>a</sup>	(.43)		(.50)		(.42)	
3. Rate of Elicitation (concepts per second)	.20	.22	.18	.20	.19	.19
Reliability <sup>a</sup>	(.21)		(.73)		(.60)	
<u>Similarity--All Concepts</u>						
4. Number of Elicited Concepts at T <sub>2</sub> Similar to T <sub>1</sub>	3.9		3.9		4.0	
5. % of Concepts Elicited at T <sub>2</sub> also Elicited at T <sub>1</sub>	48%		52%		47%	
6. Rank Order Correspondence <sup>b</sup> between Similar Concepts	.23		.09		.18	
<u>Similarity--Abstract Attributes</u>						
7. Abstract Attributes as % of Total Similar Concepts	77%		67%		71%	
8. Rank Order Correspondence <sup>b</sup> between Similar Abstract Attributes	.20		.29		.28	

Note: N = 30

<sup>a</sup>Test-retest correlation coefficient

<sup>b</sup>Mean Kendall's tau

<sup>6</sup> Due to space limitations, these data are not presented here.

<sup>7</sup> Elicited concepts were judged as concrete if they described a visual image or a sequence of images (script), or expressed a probable action, or referred to particular appropriate usage occasions. Abstract concepts were those relating to product attributes or characteristics (see Geistfeld et al., 1977).

TABLE 2  
MEANS AND RELIABILITIES FOR SELECTED MEASURES OF THE STABILITY OF  
FREE ELICITATIONS TO PRODUCT/PURCHASE SITUATION PROBES

Selected Measures	Product/Purchase Situation Probe					
	Toothpaste for own use		Ballpoint Pens for own use		Blue Jeans for own use	
	T <sub>1</sub>	T <sub>2</sub>	T <sub>1</sub>	T <sub>2</sub>	T <sub>1</sub>	T <sub>2</sub>
1. Number of Elicited Concepts Reliability <sup>a</sup>	7.1 (.76)	5.9	6.8 (.54)	6.9	8.2 (.77)	7.7
2. Time Required for Elicitation (sec.) Reliability <sup>a</sup>	47.3 (.54)	38.6	46.1 (.35)	40.3	59.1 (.52)	44.4
3. Rate of Elicitation (concepts per second) Reliability <sup>a</sup>	.16 (.32)	.16	.16 (.64)	.17	.16 (.41)	.18
<u>Similarity--All Concepts</u>						
4. Number of Elicited Concepts at T <sub>2</sub> Similar to T <sub>1</sub>	3.3		3.2		4.2	
5. % of Concepts Elicited at T <sub>2</sub> also Elicited at T <sub>1</sub>	60%		53%		56%	
6. Rank Order Correspondence <sup>b</sup> between Similar Concepts	.34		.33		.22	
<u>Similarity--Abstract Attributes</u>						
7. Abstract Attributes as % of Total Similar Concepts	63%		77%		76%	
8. Rank Order Correspondence <sup>b</sup> between Similar Abstract Attributes	.45		.28		.47	

Note: N = 30

<sup>a</sup>Test-retest correlation coefficient

<sup>b</sup>Mean Kendall's tau

TABLE 3  
MEANS AND RELIABILITIES FOR SELECTED MEASURES OF THE STABILITY OF FREE ELICITATIONS TO BRAND NAME PROBES

Selected Measures	Brands											
	Toothpaste				Ballpoint Pens				Blue Jeans			
	Crest		Ultra-Bright		Bic		Cross		Levi		Wrangler	
	T <sub>1</sub>	T <sub>2</sub>	T <sub>1</sub>	T <sub>2</sub>	T <sub>1</sub>	T <sub>2</sub>	T <sub>1</sub>	T <sub>2</sub>	T <sub>1</sub>	T <sub>2</sub>	T <sub>1</sub>	T <sub>2</sub>
1. Number of Elicited Concepts Reliability <sup>a</sup>	6.7 (.47)	6.4	5.8 (.61)	4.8	7.4 (.61)	6.2	6.1 (.79)	5.2	7.3 (.72)	5.9	6.7 (.55)	5.6
2. Time Required for Elicitation(sec.) Reliability <sup>a</sup>	41.5 (.51)	32.8	40.3 (.41)	24.9	40.0 (.40)	30.9	36.4 (.37)	24.6	44.1 (.73)	31.2	40.4 (.44)	27.6
3. Rate of Elicitation (concepts per second) Reliability <sup>a</sup>	.17 (.57)	.21	.14 (.50)	.18	.18 (.55)	.20	.17 (.57)	.21	.18 (.46)	.19	.17 (.58)	.20
<u>Similarity--All Concepts</u>												
4. Number of Elicited Concepts at T <sub>2</sub> similar to T <sub>1</sub>	3.9		2.8		3.7		3.4		3.8		3.3	
5. % of Concepts Elicited at T <sub>2</sub> also Elicited at T <sub>1</sub>	63%		64%		58%		65%		63%		62%	
6. Rank Order Correspondence <sup>b</sup> between Similar Concepts	.28		.39		.29		.24		.41		.35	
<u>Similarity--Abstract Attributes</u>												
7. Abstract Attributes as % of Total Similar Concepts	75%		72%		81%		82%		88%		92%	
8. Rank Order Correspondence <sup>b</sup> between Similar Abstract Attributes	.38		.40		.29		.29		.48		.38	

Note: N = 30

<sup>a</sup>Test-retest reliability coefficient

<sup>b</sup>Mean Kendall's tau



## Brand Name Probes

Table 3 presents the reliabilities for selected measures of the concepts elicited in response to the six brand name probes. Much the same pattern of stability and reliability was obtained as for the other probe types. The number of concepts elicited and the time and rate of elicitation were moderately reliable over the test-retest tasks. Approximately 60% of the concepts elicited at T<sub>2</sub> were also elicited at T<sub>1</sub>, and the rank order correspondence of elicited concepts was moderate (.24 to .41). Abstract product attributes constituted approximately 80% of the similar elicited concepts and the elicitation rank order correspondence among these abstract dimensions was moderately positive (.29 to .48).

## Summary and Discussion

The general pattern of results indicates that the free elicitation procedure is capable of producing responses which are moderately stable. Moreover, measures of these responses appear to be reasonably reliable. Although reliabilities of .85 to .95 might have been hoped for, such high levels of stability/reliability are unlikely for such data. The authors believe, in fact, that the levels of stability and reliability evidenced here are quite encouraging, given the nature of the elicitation task. Free elicitation responses may be influenced by a number of variables in addition to the presumed major causal factor--namely, the underlying cognitive knowledge structure in memory. Considered in this light, the obtained reliabilities and indices of stability seem acceptably strong.

Another perspective on the magnitude of the obtained reliability indices may be gained by examining the test-retest reliability coefficients obtained for variables more commonly measured in marketing research. Reliabilities for single measures of attitude and purchase intention toward each of the six brands ranged from .65 to .89. Subjects' self-reports of their past brand purchasing behavior were also quite reliable (.60 to .92). The magnitude of these reliability coefficients suggest the likely levels of reliability that can be obtained for relatively simple, direct, almost standardized measures of presumably stable constructs. Compared to these somewhat stronger levels, the reliabilities obtained for the free elicitation procedure appears "about right" and encouragingly positive, given the nature of the method and the "freedom" of response.

## Conclusions and Implications

The present results indicate that the free elicitation procedure can generate concepts that are moderately stable over repeated elicitation. Moreover, those concept responses can be described by measures that possess encouraging levels of reliability. Before examining the implications of these data, however, their limitations should be reviewed.

## Limitations

A major limitation to any test-retest study is the possibility of alternative explanations for the reliability coefficients (Silk, 1977). Like most research, this investigation unavoidably confounds the reliability of the empirical measure and the stability of the measured construct.<sup>8</sup>

<sup>8</sup> A procedure does exist which can separate the reliability of the measure from the stability of the construct being measured (Heise, 1969). However, the measure must be taken three times under a set of rather restrictive assumptions.

Thus, it is not possible to know whether, for example, a test-retest correlation of .7 means that the measure of a stable construct is moderately reliable, or that a moderately stable construct is measured by a totally reliable measure, or that the measure of a somewhat unstable construct is somewhat unreliable. In the present study, the first elicitation may have had some effects (undefined at this point) on the underlying knowledge structure in memory. If so, the present reliability coefficients reflect a certain degree of non-stability or change in the underlying construct. However, the present study did conduct the retest procedure under conditions nearly identical to the first elicitation, and therefore the possible influences on the stability of the construct were minimized. Another limitation involves the restricted external validity of these results to other subject populations, product categories, and brands. However, on purely logical grounds it would seem that the free elicitation procedure should produce equivalent levels of reliability/stability for adult consumers who have levels of verbal fluency similar to college students.

In general, the present results indicate that the free elicitation procedure can produce sufficiently stable data and reliable measures that further research with the technique is warranted. The procedure has particular implications for two issues of importance in consumer research, attribute salience and memory structure.

## Attribute Salience

Because of their concern with how consumers perceive and choose among alternative products and brands, consumer researchers are especially interested in the specific attributes or product characteristics that consumers consider salient--i.e., as "important" in some sense. Although various methods for establishing attribute salience have been used (Myers and Alpert, 1977; Sampson and Harris, 1970; Wilkie and Pessemier, 1973), the basic procedure advocated by Fishbein (1967) is an elicitation procedure. Fishbein's typical approach, however, has not been to use completely free elicitation. Rather he has given subjects rather specific instructions regarding the type of concepts desired, e.g., "tell me what you believe to be the characteristics, qualities, and attributes of. . ." (Fishbein and Ajzen, 1975; Kaplan and Fishbein, 1969). Marketers using elicitation procedures also appear to ask consumers to state "the product characteristics that come to mind. . .". Because of this emphasis on product characteristics, it is not surprising that the typical multi-attribute model questionnaire used in marketing research includes belief ratings only about relatively abstract product attributes.

It is becoming recognized, however, that other types of cognitive representations may be related to attitudes (cf. Abelson, 1976; Calder, 1978). These include concrete visual images and scripts (Shank and Abelson, 1977) as well as relatively concrete attributes (e.g., color; see Geistfeld, Sproles, and Bradenhop, 1977). Only a totally unrestricted type of elicitation task as used in this research is likely to generate these types of concepts. The present study consistently found that 20 to 30% of the stable concepts elicited on both trials were more concrete image or script type responses, while the remaining 70-80% were the more abstract product attributes. As a brief example, a common script type elicitation for the tooth-paste brand probes was the description of a scene or scenes from a familiar television commercial. Attitude theorists and applied marketing researchers should begin to consider such cognitive elements and attempt to determine their effects on attitude formation and choice behavior.

Fishbein's operationalization of salience has relied heavily on the ordering of elicited concepts, with the first few (5 to 9) supposedly most likely to be salient (or most salient). The data reported here suggest caution in interpreting elicitation orderings in this manner. The

test-retest correspondence of the elicitation rank orders were uniformly positive but not extremely strong (all tau's < .50).<sup>9</sup> If one cannot be confident of the stability of an ordering of elicited concepts, then one should not be confident in a theoretical interpretation of that order. Of course, replications using other subject samples and other products are required to determine the "stability" of this finding.

Finally, the present study found that roughly 50 to 60% of the concepts elicited in a second trial were also elicited in a previous trial. Therefore, marketers may wish to restrict their definition of salience to only those concepts that are reliably elicited over multiple probings, and not be particularly concerned about the less stable rank ordering of those similar concepts. Finally, it is interesting to note that the number of similar concepts elicited in both trials is substantially fewer (3 or 4) than the number of salient attributes typically discussed in the marketing and psychological literature (7-9 or even more; cf. Fishbein and Ajzen, 1975; Wilkie and Pessemier, 1973).

#### Memory Structure

The present results should encourage basic research into semantic knowledge structures stored in memory (Calder, 1975; Olson, in press). Data and measures derived from the free elicitation procedure which might be interpreted as indicating aspects of memory structure (e.g., the number of concepts and the time and rate of elicitation) appear to be at least moderately stable and/or reliable over time. The content of memory structure is also of interest and appears to be measurable with at least moderate reliability by the free elicitation procedure. In response to a variety of probes, a consistent 50-60% of the concepts elicited at time 2 were also elicited at time 1. In sum, free elicitation with its flexibility in terms of easy experimental manipulation of probe cues, may produce interesting insights into how consumers store and organize product/brand information.

#### Future Research

In further research with the free elicitation procedure, several issues are important. Although, to meet space constraints, the present paper did not examine the effects of individual differences factors such as product ego-involvement, it seems likely that consumers who are more highly involved with a product or brand should possess more well-developed, stable knowledge structures (Olson and Dover, 1978). Thus, such consumers should produce more stable responses (and more reliable measures) in a free elicitation task than those lower in ego-involvement. The effects of individual differences on the stability of free elicitation responses warrant research attention.

A major task for future research is to establish the construct validity of the elicited responses and the various measures thereof. Empirical evidence must be obtained that can support the interpretation of these measures as indicating the content and organization of knowledge structures in a consumer's memory. One possible approach would be to experimentally modify a memory structure, perhaps by exposing subjects to a strongly persuasive message, and then check for evidence of cognitive change by conducting a subsequent elicitation of the contents of the presumably altered structure.

Finally, there is a need to work toward further refinement of the free elicitation procedure itself. For

instance, subtle modifications in procedure may cause substantial changes in response stability. To facilitate such explorations and extensions, the present methods have been reported in complete detail.

#### Appendix

##### Instructions: Part I

We are interested in your thoughts and feeling about several common consumer products. Most of the products are ones that you have bought and probably buy fairly regularly. In the first part of our study, we are interested in your general thoughts about four products. I will tell you the name of each product, one at a time. Then, I want you to simply tell me the things that come to your mind about that product. You should concentrate on the name of the product and verbally tell me your thoughts as completely as possible. You will have about a minute-and-a-half to respond to each product. However, the time limit is not strict, so if you feel you have a lot to say, you may take more time. For some products, lots of things might come to mind. For others, perhaps only a few things will come to mind. You should stop whenever no other thoughts about the product come to mind, even if this happens before the minute-and-a-half is over. But, please tell me all the thoughts you have about each product. In giving your responses, you may have to pause to form your thoughts. Don't let that pause bother you. It is only natural that you might need a few seconds for certain thoughts. But when you do think of something, please state it right away. Please do not censor your comments. We are interested in what you think. If something comes to mind, please say it--even if you might consider it unimportant, ordinary, or even silly. In giving your thoughts, please state each one in short phrase or a word or two. If you feel that you must use more words to express yourself completely or clearly, please feel free to do so. However, we are not interested in why you think your responses are important. So, please do not give justifications or elaborate descriptions of your thoughts--just a simple statement of each thought is sufficient. You should try to give your thoughts as rapidly as possible. Try not to let me distract you in any way. I am just here to give you the instructions and the product names. You might try, either psychologically, or by physically turning in a different direction, or by closing your eyes, to disregard me. I will be recording both our comments during this session on this tape recorder. I am taping the sessions only to be sure I get an accurate record of your responses. However, your name will never be associated with the recording or your specific responses and, when we are finished with the study, the tapes will be erased. Before we start, let me summarize the procedure again. I will first tell you the name of the product. Then you tell me all the things that come to your mind about this product, as rapidly and completely as possible. Do you have any questions before we begin?

##### Instructions: Part II

In this part of the study, we are interested in your thoughts about each of these products when you are considering buying the product in a specific situation. First, I will explain to you the buying situation of interest. For example, you might assume that you are thinking of buying this product for a specific purpose or occasion. Then, I will tell you the name of the product. I want you to tell me the things that come to your mind about buying that product in that situation. You should try to concentrate on the product and the situation, together, and give me your thoughts as rapidly and completely as possible, just as before.

<sup>9</sup> Interpretation of the low positive rank order correlations must be tempered by the recognition that when the number of concepts is few, as in the present case (3 or 4 similar concepts was typical), a single reversal in rank order dramatically lowers the tau coefficient.

Instructions: Part III

In this part of the study we are interested in your thoughts about particular brands of certain products. First, I will tell you the name of a brand of one of the products we have been concerned with today. Then, you tell me the things that come to your mind about that brand. Concentrate on the brand and give me your thoughts as rapidly and completely as possible, just as before.

References

- Robert P. Abelson, "The Structure of Belief Systems," in R. C. Schank and K. Colby, eds., Computer Models of Thought and Language (San Francisco: W. H. Freeman, 1973)
- 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.: Lawrence Erlbaum, 1976)
- Herman Buschke, "Two-Dimensional Recall: Immediate Identification of Clusters in Episodic and Semantic Memory," Journal of Verbal Learning and Verbal Behavior, 16 (April, 1977), 201-15.
- Bobby J. Calder, "The Cognitive Foundations of Attitudes: Some Implications for Multi-Attribute Models," in M. J. Schlinger, ed., Advances in Consumer Research, Vol. 2 (Chicago: Association for Consumer Research, 1975)
- Bobby J. Calder, "Cognitive Response, Imagery, and Scripts: What is the Cognitive Basis of Attitude?" in H. Keith Hunt, ed., Advances in Consumer Research, Vol. 5 (Ann Arbor, Mich.: Association for Consumer Research, 1978)
- 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), 456-60.
- George S. Day, "Evaluating Models of Attitude Structure," Journal of Marketing Research, 9 (August, 1972), 279-86.
- Martin Fishbein, "A Behavior Theory Approach to the Relationship Between Beliefs About an Object and the Attitude Toward the Object," 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)
- W. A. K. Frost and R. L. Baine, "The Application of the Repertory Grid to Problems in Marketing Research," Commentary, 9 (October, 1967), 161-75.
- Loren V. Geistfeld, George B. Sproles and Suzanne B. Bradenhop, "The Concept and Measurement of a Hierarchy of Product Characteristics," in W. D. Perreault, Jr., ed., Advances in Consumer Research, Vol. 4 (Atlanta: Association for Consumer Research, 1977)
- Paul E. Green, Yoram Wind and Arun K. Jain, "Analyzing Free Response Data in Marketing Research," Journal of Marketing Research, 19 (February, 1973), 45-52.
- David R. Heise, "Separating Reliability and Stability in Test-Retest Correlation," American Sociological Review, 34 (February, 1969), 93-101.
- Eric J. Johnson and J. Edward Russo, "The Organization of Product Information in Memory Identified by Recall Times," in H. Keith Hunt, ed., Advances in Consumer Research, Vol. 5 (Ann Arbor, Mich.: Association for Consumer Research, 1978)
- Kalman J. Kaplan and Martin Fishbein, "The Sources of Beliefs, Their Saliency, and Prediction of Attitude," Journal of Social Psychology, 78 (June, 1969), 63-74.
- George A. Kelly, The Psychology of Personal Constructs (New York: Norton, 1955).
- O. Hobart Mower, "Mental Imagery: An Indispensable Psychological Concept," Journal of Mental Imagery, 1 (Fall, 1977), 303-26.
- James H. Myers and Mark I. Alpert, "Determinant Buying Attributes: Meaning and Measurement," Journal of Marketing, 32 (October, 1968), 13-20.
- 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 (Atlanta: Association for Consumer Research, 1977)
- Jum D. Nunnally, Psychometric Theory (New York: McGraw-Hill, 1967)
- Jerry C. Olson, "Theories of Information Encoding and Storage: Implications for Consumer Research," in Andrew A. Mitchell, ed., Effects of Information on Consumer and Marketing Behavior (Chicago: American Marketing Association, 1978, in press)
- Jerry C. Olson and Philip A. Dover, "Attitude Maturation: Changes in Related Belief Structures Over Time," in H. Keith Hunt, ed., Advances in Consumer Research, Vol. 5 (Ann Arbor, Mich.: Association for Consumer Research, 1978)
- J. Paul Peter, "Reliability: A Review of Psychometric Basics and Recent Marketing Practices," Unpublished paper, Washington University, 1977.
- Peter Sampson, "Using the Repertory Grid Test," Journal of Marketing Research, 9 (February, 1972), 78-81.
- Peter Sampson and Paul Harris, "A User's Guide to Fishbein," Journal of the Market Research Society, 12 (July, 1970), 145-89.
- Roger Schank and Robert Abelson, Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structures (Hillsdale, N.J.: Lawrence Erlbaum, 1977)
- Carolyn W. Sherif, Muzafer Sherif and R. E. Nebergall, Attitude and Attitude Change: The Social Judgment Involvement Approach (Philadelphia, Pa.: W. B. Saunders, 1965)
- Alvin J. Silk, "Test-Retest Correlations and the Reliability of Copy Testing," Journal of Marketing Research, 14 (November, 1977), 476-86.
- William L. Wilkie and Edgar A. Pessemier, "Issues in Marketing's Use of Multi-Attribute Attitude Models," Journal of Marketing Research, 10 (November, 1973), 428-41.
- David T. Wilson and Philip A. Dover, "A Test of the Repertory Grid Technique as a Means of Developing Attributes for Use in Choice Models." Paper No. 34, Working Series in Marketing Research, College of Business Administration, Pennsylvania State University, 1975

DISCUSSION COMMENTS: "ATTRIBUTES: IMPORTANT, SALIENT, AND IDEAL"

Richard P. Bagozzi, University of California at Berkeley

Introduction

The three articles presented in this session all, in one way or another, address the issue of the importance of product attributes for consumers. Despite this commonality in overall purpose, however, each article uses a vastly different methodology, and each finds different results with different implications for consumer research. The comments below represent an attempt to point out possible shortcomings of the procedures and certain conceptual limitations. Effort is made also to provide suggestions for improving the research and its interpretation. Nevertheless, the reader should note that the articles represent pioneering efforts for assessing attributes, and their strengths far outweigh any limitations that may be signalled-out.

Behavioral Measurement of the ...  
by John A. Quelch

This is a thorough, well thought-out piece of research dealing with an important area in consumer information processing: i.e., the relative importance of attributes in the choice process. There are at least two broad contributions of the research that deserve mention. First, the amount of methodological control and effort taken to discuss the meaningfulness of the research design and general problem are exemplary. The author has operationalized the information display methodology in a manner achieving a good balance between internal and external validity. Overall, the methodology attains a reasonable degree of control of stimuli and external factors while introducing a certain degree of realism in the study. A second contribution can be found in the findings. Specifically, the results provide an interesting descriptive picture of the acquisition of information in terms of the number of attribute dimensions used and the order of information elicitation. Further, the effect of providing a physical sample on information use is ascertained.

Because the strengths of the research can be readily seen on a reading of the article, I will focus on a number of shortcomings and limitations in the study. Further, comments will be made on ways to improve the study or better address the general problem area.

A first question to ask is whether the study can really be viewed as one investigating information processing in the context of initial trial and repeat buying. Although all subjects were recent purchasers and consumers of cold breakfast cereal, a possible confound or artifact exists with the second treatment group. This group--which was exposed to the physical product--was told that the brands were not available in their area. Yet, because the brands were, in fact, real, well-known, and leading sellers in the respondents market, it is likely that many would recognize the brands or at least some of them. This would most likely occur differentially across people in the second sample. Moreover, given the fact that the stimuli are so easily recognized (e.g., Corn Flakes, Rice Chex, Raisan Bran), it is also likely that some of the respondents tried to guess the manufacturer or brand or drew parallels to their own consumption experiences. Both of the above factors could introduce bias in the sense that the inferences made would contain unintended confounding information. At a minimum, a manipulation check could have been performed to see if the respondents received contaminated information and believed the ruse. Further, it seems

incorrect to label this sample the repeat purchase condition. What evidence is there that the subjects all regarded the second stimulus as a familiar one?

A second problem entails a possible demand characteristic. Given the particular information display design, it may have been the case that respondents were forced to adopt a lexicographic-like strategy in such a way as to conceive of the experimental task as a game. That is, they may have viewed the task as requiring that they "guess the best or favorite brand using the least amount of information." Notice that it is possible that the "best" brand need not be best relative to their own criteria for decision making but rather might reflect what they think the experimenter, their family, or society expects of them. Evaluation apprehension and demand characteristics seem likely and should have been investigated, perhaps with a manipulation check. Notice further that had the respondents perceived the task as entailing choosing a brand using the least amount of information possible (a reasonable normative expectation people might form when asked to perform in an experiment), then their information processing might reflect a tendency to select and focus on cues offering the greatest potential for a fast judgment. These cues may or may not be the most important to them with respect to consumption of cereal for their own use. Thus, both the internal and external validity of the study may be less than desirable.

The data might support a rival hypothesis in still another sense. Rather than being an indicator of relative importance, the rank order elicitation might merely reflect an "ease of information processing factor." In particular, people might evaluate alternatives by scrutinizing the easiest, least costly, or least time consuming attributes first. The ease of evaluation or cost may not at all be related to the importances of the attributes.

One of the assumptions stated in the study was that the most important attributes would be elicited early "with a view to minimizing the duration of the decision-making process." Alternatively, if the respondent's decision rule were one of accuracy or minimizing risk, it might be expected that the most important attributes would be selected first but with time not considered to be a salient consideration. In any case, the rationale for why the most important attributes would be selected in any order has multiple interpretations and needs further development.

Closely related to this point is the role played by the compensation. Ostensibly, the coupons were intended to motivate the subjects to behave in a manner similar to natural purchase occasions. However, it can just as easily be argued that the coupons motivated subjects to choose unknown or new information and products. That is, some subjects might have viewed the coupons as an incentive to take a chance on a new product--to see if it is any good. The induced motivation may have been also one of trying a product that is new because it costs less or is a "free good." In addition, when looking across subjects within each treatment, decision-making differences might have been a function of individual variance in venturesomeness, habit, or other factors. Hence, conclusions made across subjects within treatments must be regarded with suspicion.

Other questions to answer are: What is the construct validity of the importance measures? Were the differences in decision making across treatment groups statistically significant? Is the experiment a test of the relative importance of attributes? It seems that it is really a test of classes of information or bundles of undifferentiated attributes. Moreover, the design appears to test only the effect of "the opportunity to elicit physical samples," since this is the only difference between the two treatment groups. What does such a manipulation mean and how does "attributes" conceptualized in this way relate to past research on the role of attributes in decision making?

To place things in a proper perspective, a number of comments deserve mention. First, the study deserves recognition as an innovative piece of research. Overall, it was handled with considerable care and foresight. Second, despite the criticisms that can be made, the study does achieve a certain degree of internal validity with some realism provided through the operationalization of stimuli and the general design. The shortcomings that have been pointed out are primarily a function of the attempt to achieve both internal and external validity in a single study. Such a goal will invariably involve tradeoffs and produce pros and cons as part of its very nature. Finally, it should be noted that the article by Professor Quelch represents a well-organized and articulated investigation that should be read by anyone seriously interested in conducting research into consumer decision making.

A Comparison of Two Methods...  
by James H. Myers

The problem of determining the optimal level of product attributes is an important one, and Professor Myers presents an interesting study comparing two alternative methodologies. Perhaps the most significant issue in the research to note is the finding that the direct method for determining the ideal point (which is based on subjective ratings) provides consistently lower values than the indirect method (which is based on a cross-tabulation of beliefs versus evaluations). The question arises as to whether one method is more accurate than the other, or for that matter, are either valid procedures?

The question of accuracy or validity is not an easy one to address. Because the research begins with measurements of concepts at the level of individual behavior and then aggregates to the group level, it is important to scrutinize the entire process of measurement and aggregation. For example, consider the method for determining direct ideal points. The respondents were asked to rate their ideal using a six-point scale on each of eleven attributes for each of four products. For each attribute, an ideal point score was determined by first obtaining an average for each individual (based on the sum over the four products) and then averaging across people. The resulting score is thus an ideal rating for the attribute based on all respondents. What is the construct validity of this score? Notice that measurement error occurs at least in two places: (1) at the point of averaging across products and (2) at the point of averaging across respondents. Notice further that whatever measurement error is present at the first stage is added to or confounded with that found at the second. Because measurement error is not taken into account or modeled explicitly, it is possible that the ideal point construct is highly fallible. In any case, given the procedure, it is not possible to assess construct validity.

One way to approach the question of construct validity might be the following. First, the validity of the ideal point model could be evaluated at the level of the individual responses. Three components of construct

validity need to be addressed: namely, convergent, discriminant, and nomological validity (for a model and methodology for determining construct validity, see Bagozzi (1977, 1978a, 1978b)). The assessment of convergent and discriminant validity requires that multiple measures and methods be employed. The evaluation of nomological validity requires that another concept be employed in a predictive context along with the measures of the ideal point. Second, after establishing the accuracy of measurements and validity of the ideal point ratings at the individual level, the scores can then be aggregated to form a distribution for the market conditions. It is important that the sample of individuals used to summarize the aggregate be a representative one, perhaps drawn randomly from the potential market. Notice that the results obtained by either of the two methods used by Professor Myers are directly dependent on the quality of the sample chosen. In addition to the internal consistency of ideal point measures at the individual level, the validity of the optimal level of attributes obtained in the aggregate depends also on the representativeness of the distribution of responses across people and ratings.

One further problem with the procedures used by Professor Myers deserves mention. The data are derived by observing responses to four products which have fixed values of the eleven attributes. A better procedure would be to obtain descriptive and evaluative ratings on each attribute for each of a number of levels. For example, the response of each person to each of five levels of saltiness could be obtained. Although this increases the time for data collection considerably, some time can be saved if only five or six attributes are investigated. Most products probably possess only five or six salient attributes anyway rather than eleven as used in the present study. Alternatively, an experimental design with blocks, such as the Latin square, can be used to achieve the desired economies.

Finally, it should be noted that the derived methodology may not go far enough in that the cross-tabulations may overlook complex interactions among attributes. For example, no relationship may be evident between characteristic B and evaluations when viewed in isolation, but a strong relationship might exist when characteristic E is taken into account. To represent interaction effects, a multivariate analysis is required such as provided by the log-linear model. For instance, a model to test for all interactions among three salient attributes (A, B, and C) would be

$$L_{ijk} = \mu + \mu_i^A + \mu_j^B + \mu_k^C + \mu_{ij}^{AB} + \mu_{jk}^{BC} + \mu_{ik}^{AC} + \mu_{ijk}^{ABC}$$

where L is the natural logarithm of expected frequencies and the  $\mu$ 's represent main and interaction effects analogous to the terms in the ANOVA model. A procedure such as Goodman's (1970) ECTA methodology can be used to estimate parameters and test hypotheses. The derived methodology rests on the degree of association between attributes and evaluations. Thus, these relationships must be ascertained before confidence in any overall rating can be gained.

Finally, the nature of the relationship between attribute belief judgments and evaluations needs to be examined. In the cross-section survey, the relationship may be purely correlational, causal, or spurious. Yet, the cross-tab cannot identify which of these alternatives it is. An experimental methodology can at least ascertain whether the relationship is a causal one from beliefs to evaluation. The latter is necessary if management is to have faith in its research and decisions to fix product attribute levels.

This article represents an important contribution for at least two reasons. First, the operationalization of the free elicitation procedure itself is significant in that it demonstrates that such a method can yield meaningful data on cognitive responses. Because it does this in a relatively unobtrusive and unrestrictive way, it provides a promising means for measuring a variety of cognitive concepts central to contemporary theories of human decision making. Second, the test-retest results suggest that the overall methodology may yield reliable measures for at least some concepts of interest in the free elicitation context. With these strengths of the study in mind, the following comments will focus on a number of caveats and suggestions for extending the research.

A first issue to note is the meaning of the concept of stability. The authors state that "stability means that a probe cue...presented to consumers at two points in time elicited similar responses" (as reflected by the rest-retest correlation). It is important to note that stability as the authors define and use it is a "floating" measure of change in that it depicts the change in each individual respondent's position in the distribution of responses *relative* to other individuals. Thus, interest is restricted to the slope of responses between time periods rather than a change in intercept which might reflect an absolute, main-effect change in the group as a whole. The authors' definition of stability is an important one, but it should be mentioned that it does not address stability in the sense of a systematic influence on responses experienced by all respondents between measurements. Rather, it in effect assumes that no such shock occurred to any or all individuals. If some respondents experienced an extraneous influence between the times of measurement, then this would contaminate the test-retest correlation. Given the present research, there is no way to assess the possibility of this one way or another.

The use of the test-retest correlation as a measure of stability is fine as far as it goes but a number of limitations of the procedure deserve mention. First, because situational factors may systematically affect people, the correlation between measures could be a function of both the characteristics of people (for which the measurement is intended to capture) and the external influence. Yet, the usual test-retest procedure cannot separate these different factors. Second, because the cognitive states of people change over time due to fatigue, maturation, learning, differences in motivation, and so on, the error in measurement from one time to another can also change systematically, producing biased results. The test-retest correlation does not allow one to discriminate between such effects and the true stability over time. Third, demand characteristics, evaluation apprehension, and memory effects tend to be present at each measurement occasion, causing a correlation between errors in measurement and a corresponding bias in the observed correlation between measures.

One way to better represent stability--given that one is restricted to single operationalizations of each concept at each point in time--is to employ the simplex model or a variation thereof. Although the model suggested by Heise (1969) makes rather restrictive assumptions as the authors note, the one proposed by Wiley and Wiley (1970) is somewhat less restrictive. Both models require that measures be taken at three points in time. The Heise (1969) and Wiley and Wiley (1970) models begin with the following structure:

$$y_1 = \alpha_1 T_1 + u_1$$

$$y_2 = \alpha_2 T_2 + u_2$$

$$y_3 = \alpha_3 T_3 + u_3$$

$$T_2 = \beta_1 T_1 + \theta_1$$

$$T_3 = \beta_2 T_2 + \theta_2$$

where the  $y_i$  are measures of a relevant concept; the  $T_i$  are the corresponding true scores; the  $\alpha_i$  and  $\beta_j$  are parameters relating measures to true scores and true scores to past true scores, respectively; and each error term is mutually uncorrelated with all other error terms and the explanatory variables. Given this model and the following assumptions, it is possible to calculate the reliability of the concept under consideration. With  $\alpha_1 = \alpha_2 = \alpha_3$  imposed as a constraint, Heise (1969) shows that it is possible to compute the stability parameters,  $\beta_1$  and  $\beta_2$ . In effect, this assumes that the contemporaneous reliability of each concept is equal at each point in time (although stability may change). Wiley and Wiley (1970) demonstrate that one need not assume that  $\alpha_1 = \alpha_2 = \alpha_3$  if it is assumed that the error variances are homogeneous (i.e.,  $\text{var}(u_1) = \text{var}(u_2) = \text{var}(u_3)$ ). Werts, Jöreskog, and Linn (1971) further demonstrate that neither homogeneity in contemporaneous reliabilities nor homogeneity in error variances is necessary to assume, if one has measurements at four points in time. Their model, however, does constrain the contemporaneous reliability parameters to equal one. Further, their model generalizes to multiple wave panels with the caveat that "error variances, true score variances, and unstandardized regression weights between corresponding true scores are identified for all but the first and last measures" (Werts, et al., 1970, p. 111). The procedures developed by Werts, et al. offer the advantages over the traditional test-retest correlations of explicitly modeling measurement error and representing stability as an association between true scores.

If multiple measures are available for a variable of interest for at least two points in time, then the general analysis of covariance structures approach can be used to assess concurrent reliability and stability simultaneously (cf., Bagozzi, 1978b, c). For example, given two measures of a variable at two points in time, the structural equation model permitting the determination of concurrent reliability and stability can be written as:

$$\tilde{x} = \tilde{A}\tilde{T} + \tilde{u}$$

$$\begin{pmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \end{pmatrix} = \begin{pmatrix} \alpha_{11} & 0 \\ \alpha_{12} & 0 \\ 0 & \alpha_{21} \\ 0 & \alpha_{22} \end{pmatrix} \begin{pmatrix} T_1 \\ T_2 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix}$$

$$\tilde{\Sigma} = \tilde{A}\tilde{\Phi}\tilde{A}' + \tilde{\Psi}$$

where it is assumed  $E(x) = E(T) = E(u) = 0$ ,  $E(TT') = \Phi$ ,  $E(uu') = \Psi$  (a diagonal matrix of error variances), and  $E(xx') = \Sigma$ . Using a maximum-likelihood estimation procedure, the stability coefficient may be estimated as the appropriate covariance in  $\Phi$ . Similarly, contemporaneous estimates of the individual and composite reliabilities can be computed from the following formulae, respectively:

$$\rho_{x_i} = \frac{\alpha_{x_i}^2 \text{var}(T_j)}{\alpha_{x_i}^2 \text{var}(T_j) + \text{var}(u_{x_i})}$$

$$\rho_c = \frac{\left( \sum_{i=1}^{n_j} \alpha_{x_i} \right)^2 \text{var}(T_j)}{\left( \sum_{i=1}^{n_j} \alpha_{x_i} \right)^2 \text{var}(T_j) + \sum_{i=1}^{n_j} \text{var}(u_{x_i})}$$

where  $n$  = the number of measures at time  $j$ . For a derivation of the above procedures and an illustration, see Bagozzi (1978b, c). The stability coefficient yielded by the procedures is analogous to those corrected for attenuation. Further, the confounding of measurement error is avoided.

One final point to note is that the stability of responses derived from a free elicitation procedure can be no more meaningful than the validity and significance of the concepts measured. While such notions as the "number of elicited concepts," "time required for elicitation," and "rate of elicitation" are useful as far as they go, it would be much more meaningful to measure concepts with more substantive content. For instance, in the area of cognitive research, concepts such as the nature and intensity of counterarguments, source derogations, or support arguments might be more interesting than merely their rates or frequencies.

#### References

- Richard P. Bagozzi, "Convergent and Discriminant Validity by Analysis of Covariances Structures: The Case of the Affective, Behavioral, and Cognitive Components of Attitude," in W. D. Perreault, Jr., (ed.), *Advances in Consumer Research*, Vol. IV, (Atlanta: Association for Consumer Research, 1977), 11-18.
- Richard P. Bagozzi, "The Construct Validity of the Affective, Behavioral, and Cognitive Components of Attitude by Analysis of Covariance Structures," *Multivariate Behavioral Research*, 13 (January 1978a), 9-13.
- Richard P. Bagozzi, "Reliability, Validity, and Theory Testing in Marketing Research: A Structural Equation Approach." Unpublished paper, University of California, Berkeley, 1978b.
- Richard P. Bagozzi, "Reliability Assessment by Analysis of Covariance Structures," in S. C. Jain, (ed.), *Research Frontiers in Marketing: Dialogues and Directions*, (Chicago: American Marketing Association, 1978c), 71-75.
- Leo A. Goodman, "The Multivariate Analysis of Qualitative Data," *Journal of the American Statistical Association*, 65 (1970), 226-56.
- D. R. Heise, "Separating Reliability and Stability in Test-retest Correlation," *American Sociological Review*, 34 (1969), 93-101.
- C. E. Werts, K. G. Jöreskog, and R. L. Linn, "Comment on 'The Estimation of Measurement Error in Panel Data,'" *American Sociological Review*, 36 (1971), 110-113.
- D. E. Wiley and J. A. Wiley, "The Estimation of Measurement Error in Panel Data," *American Sociological Review*, 35 (1970), 112-117.

## SEARCHING FOR IMPORTANT ATTRIBUTES AND APPROPRIATE LEVELS

Donald R. Lehmann, Columbia University

The concept of preference and subsequently choice resulting from how alternatives are perceived on a set of key attributes is well established (Fishbein, 1967; Lancaster, 1966; Rosenberg, 1956; Wilkie & Pessemier, 1973). Work on such models requires identification of the key attributes and the desirable levels of these attributes. This paper will examine the recent contributions made by Myers (1978), Olson and Mudderisoglu (1978), and Quelch (1978) in this area, as well as some problems with their results. In addition, some other issues will be discussed, including whether attribute-based decision making is a sound model of actual behavior and if so, what methods might be used to study it.

Assuming we are interested in studying attribute-based decisions, the issue of for what purpose becomes relevant. Some of the major choices which must be made include:

### 1. Prediction vs. Understanding of Process.

Models based on multiattribute models usually work well predictively. When the issue is process understanding or causality, however, much greater care is needed in establishing their usefulness.

### 2. Group vs. Individual Behavior.

For providing guidance to product designers or managers, group level modeling is often sufficient. Studying process, on the other hand, leads to a focus on individual behavior.

### 3. Subject of Investigation.

Studying key attributes has at least three definable categories: determination of what are the key attributes, determination of the relative importances of the key attributes, and determination of optimum levels on the attributes.

Here the papers by Olson and Mudderisoglu and Quelch focus on individual behavior and process while the paper by Myers is oriented toward management decisions and aggregate behavior.

Olson and Mudderisoglu

Olson and Mudderisoglu study the determination of key attributes by means of free elicitation. Their major concern is in determining how reliable such free elicitation producers are under three different cues: product category, product/purchase situation, and brand.

### Strengths

#### 1. Important Topic.

The issue of the stability of free elicitation procedures in specific and attribute importance determination procedures in general is crucial. Yet, few authors have explicitly set out to determine how stable such responses are.

#### 2. Execution.

The procedures followed seem carefully executed, something which is lacking in many studies.

### Problems

#### 1. The Sample.

The usual caveats about a sample of 30 students are relevant.

#### 2. Procedure.

The procedure itself, though carefully executed, has some problems. First, the one-week time period between measurements may have been too short to avoid carryover effects over time (e.g., either remembering answers given in a previous week or being bored with the procedure the second time). Since the products were in stable categories (and hence few real changes would be expected), a longer period would be somewhat more desirable. Also, by using three cues sequentially, there is a real possibility of carryover across the three cues. An analysis of the similarity of the elicitations generated across the three cues would have been interesting in that regard. Finally, it would have been very interesting to see how the stability of free elicitation responses compared with that of other methods, but obviously this extension would be a major one.

#### 3. Results.

There may be more in the data than is discussed. For example, it is interesting that the number of elicitations seems to be in the  $7 \pm 2$  range. Similarly, the apparent decrease in the number of elicitations and even more dramatically in the time taken by subjects might be attributable to learning plus a wearout in the novelty of participating in the study. Also, it would be interesting to know what percent of the responses appear to be attributable to advertising playback.

### Summary

Olson and Mudderisoglu have presented some interesting data concerning the stability of free elicitation procedures. Assuming this objective, accounting for individual differences (as they suggest) as well as extensions to more complex product categories seems appropriate. More broadly, comparison of both the results and the stability of free elicitation with other procedures seems appropriate future research.

Quelch

Quelch uses an information display board methodology to uncover the importance of a variety of attributes about cold cereals.

### Strengths

#### 1. Sample.

Using 250 real shoppers is commendable.

#### 2. The Concept.

The concept of relating attribute importance to information acquisition is appealing.

### Problems

#### 1. The Procedure.

The entire procedure seems artificial. There are only five dimensions: price and two natural pairs; ingredient and nutritional information and physical appearance and physical product. These dimensions are compound, thus muddying the interpretation. Forcing respondents to obtain information about all six brands on the first attribute chosen seems unnatural. One also wonders if the subjects were able to deduce the



name of the products (e.g., Kellogg's Corn Flakes) from the appearance.

2. The Repeat Purchase Issue. The paper begins by mentioning trial and repeat choice. The data, however, is a single choice of an unbranded product and hence all the choices are really trials.

3. The Determinance Measure. The determinance measure may be flawed. It is defined as the number of brands selected on a given dimension minus the number selected on the previous dimension. This suggests that people always look for fewer brands as they proceed (which they do not always do unless constrained to). Also the measure is at least potentially order dependent (the first attribute has a potential score of 6, the subsequent ones do not) and a determinance measure based on the percent of the remaining brands eliminated rather than the raw number seems preferable.

4. The Relation of the Measures of Importance. Three measures of importance based on information acquisition are suggested and a separate constant sum scale measure is available, yet the resulting average importances are not presented. Moreover, one has no idea how closely the different measures are related to each other on an individual basis. Also comparison with other results (e.g., Jacoby, Szybillo, and Busato-Schach, 1977) would have been useful.

#### Summary

This pilot study attempts to deduce attribute importances from information display boards. As such it provides an interesting starting point for research in this area, but one which needs further refinement.

#### Myers

Myers does not directly assess the importance of attributes. Rather he attempts to provide guidance to product designers. In doing so he focuses on finding the ideal levels of descriptive attributes which can be used in guiding laboratory personnel in product design.

#### Strengths

1. The Sample. This paper has a sample size of 621 real people rating products they actually use.

2. Cleverness. The approach is clever, simple, and easy to use. It allows anyone with a knowledge of cross-tabs to indirectly deduce ideal levels on a set of attributes.

3. Usefulness. Assuming the results of this approach were correct, they are directly useful in solving a real problem, product design.

#### Problems

1. The Attributes Used. The attributes used (by being not evaluative) may also not be relevant to the consumer and hence the resulting product may be of limited appeal. In short, the whole issue of determinant attributes becomes relevant (Myers and Alpert, 1968; Alpert, 1971).

2. The Causality Assumption. The method implicitly assumes the direction of causation is from attribute position to attitude. Yet there is considerable evidence (e.g., Beckwith and Lehmann, 1975) that causation, at least in measured responses, runs both ways. Also measured ratings on any dimension will generally have an evaluative component may be especially large on less important attributes.

3. Collinearity. Examining the attributes one at a time assumes that the attributes are independent, which they almost certainly are not, and hence the estimates are biased. The best solution to this problem is a multivariate procedure which simultaneously estimates the effects of different levels of the attributes via ANOVA or dummy variable multiple regression.

4. Level of Aggregation. As in almost any model, the question of the appropriateness of aggregating people is relevant. Here people were grouped by favorite brand, which is a step in the right direction. Still the question of segment homogeneity could have been more fully explored.

#### Summary

Myers has presented an interesting approach for deducing the ideal level of attributes. If the procedure is extended to be multivariate (which is easy to do) and the direction of causality can be established more unambiguously (which is hard to do), then the procedure can be useful. Incidentally, the procedure can easily be adapted to estimating attribute importances. The range of the mean scores in the cross-tab approach (or dummy variable coefficients in the preferred multivariate approach) can be treated as a measure of importance as in conjoint analysis (Green and Wind, 1975).

#### Considerations for Future Research

##### Are Attributes Relevant

All of the papers here and much of the research in marketing, including that of the author, is based on the assumption that consumers make trade-offs on attributes. Yet this fundamental tenet deserves re-examination.

The view that consumers are actively engaged in evaluating alternatives on attributes is clearly an appealing and flattering view of man. Unfortunately, at least for most decisions it is probably also false. Every day is filled with enough decisions so that only a few can consciously be considered in anything like the fullness that multiattribute models imply. The rest are handled by rules or standard operating procedures which have been developed over time. Hence when consumers give responses about which attributes are relevant, the responses may be as much rationalization as truth. While simplified rules make more sense than complex ones (Wright, 1975), even simple ones may be employed infrequently.

Putting the proposition differently, multiattribute models are probably a good descriptive or predictive model of how consumers behave when they are heavily involved in a decision. (They may or may not accurately reflect the actual process followed by consumers.) Similarly in a market where information is readily available and the marketplace relatively free, multiattribute models are likely to fairly accurately forecast the equilibrium share of competing brands. For most situations, however, it is inappropriate to assume that individuals follow the process of information acquisition and processing implied by multiattribute models. (The situations under which such models are likely to be used are described in Lehmann, 1978.)

The importance of this is that responses to probes about what attributes are being used may measure the attributes which were used in developing a decision rule. Alternatively the responses may reflect convenient attributes to talk about, attributes which "should" be important (e.g., nutrition), or advertising playback. In summary, then, it is advisable to ask the question, "Are the respondents likely to be consciously using attributes at all?"

Alternative Methodologies

Assuming consumers can reasonably be expected to be using multiattribute models, a large variety of methodologies are available for studying the process (Table 1), all of which measure somewhat different things and which have fairly well known strengths and weaknesses. The major weakness of all of them is that they cannot directly measure what is going on inside a consumer's head. The methodologies can be classified in terms of direct or indirect in their data collection strategies. Methods for asking consumers to directly provide information include the following:

1. Elicitation can be used for determination of key attributes or importance weights, but is not useful for determining optimum levels. This method tends to measure conscious associations as recalled in memory and is affected by quality of memory, verbal ability, etc.
2. Adjective Check Lists are useful for reducing a long list of key attributes to a smaller one, and can be used to determine attribute importances.
3. Paired Comparisons of attributes can be used to determine the relative importance of pre-specified attributes, especially when graded paired judgments are collected.
4. Scaled Importance Ratings are commonly used to rate the importance of a list of attributes on a scale (e.g., 1 to 6).
5. Constant Sum Scales can also be used to get the relative importance of key attributes.

Indirect data collection procedures include:

1. Information Boards measure acquisition and hence can be used to indicate which attributes are most often used (Jacoby et al 1977). Whether this measures importance, curiosity, or failure of memory is unclear, however.
2. Eye Movements measure which attributes are being examined and hence also may indicate importance (Russo, 1977), although some obvious problems exist with this approach as well as other physiological measures.
3. Protocols can be used to elicit both which attributes are being used and how important they are. Recall protocols, however, suffer many of the problems of elicitation procedures.
4. Brand Rating and Preference Data can be used to deduce attribute importance or optimum levels by using regression analysis (Beckwith and Lehmann, 1978) or programming methods (Srinivasan and Shocker, 1973; Pekelman and Sen, 1974).
5. Multidimensional Scaling can be used to deduce both the content and importance of attributes as well as their ideal levels. However, similarity based maps may not reflect preference dimensions and the procedure is limited to a small number of attributes whose identity is always uncertain.
6. Conjoint Analysis can be used with a small number of dimensions to find both desired levels and relative importance.

Conclusion

Searching for key attributes and attempting to measure their relative importance and most desired levels is an important area of research. However, in many instances individuals may not be consciously using attributes in

making decisions. In such cases, attempts to determine key attributes can be deceptive.

Assuming multiattribute models are appropriate or useful measures of behavior, a variety of methods are available. Which method is used depends to a large extent on the objective of the analysis. However, it seems clear that a research strategy which employs several of the methods stands a better chance of overcoming their individual weaknesses than a monolithic approach to the problem.

TABLE I  
METHODOLOGIES AVAILABLE FOR STUDYING KEY ATTRIBUTES

Methodology	Subject of Investigation		
	Determination of Key Attributes	Determination of Importance Weights	Determination of Desired Levels
I. Direct Data Collection			
A. Elicitation	x		
B. Adjective Check List	x		
C. Paired Comparisons			x
D. Scaled Importance Ratings			x
E. Constant Sum Scale			x
II. Indirect Assessment			
A. Information Acquisition			
1. Information Boards			x
2. Eye Movements			x
3. Protocols	x		x
B. Brand Attribute Ratings Plus Preference Based: (ANOVA, Regression, Programming)			x
C. Similarity Based: MDS	x		x
D. Conjoint Analysis		x	x

REFERENCES

Mark I. Alpert, "Identification of Determinant Attributes: A Comparison of Methods," Journal of Marketing Research, Vol. 8, May, 1971, pp. 184-191.

Neil E. Beckwith and Donald R. Lehman, "The Importance of Differential Weights in Multiple Attribute Models of Consumer Attitude," Journal of Marketing Research, Vol. 10, May, 1973, pp. 141-145.

Neil E. Beckwith and Donald R. Lehmann, "The Importance of Halo Effects in Multi-Attribute Attitude Models," Journal of Marketing Research, Vol. 12, August, 1975, pp. 265-275.

Fishbein, Martin, "Attitude and the Prediction of Behavior," in Martin Fishbein, ed., Readings in Attitude Theory and Measurement New York, John Wiley & Sons, 1967, pp. 477-492.

Paul E. Green and Yoram Wind, "New Way to Measure Consumers' Judgments," Harvard Business Review, Vol. 53, July-August, 1975, pp. 107-117.

Jacob Jacoby, George J. Szybillo, and Jacqueline Busato-Schach, "Information Acquisition Behavior in Brand Choice Situations," Journal of Consumer Research, Vol. 3, No. 4, March, 1977, pp. 209-216.

Jacob Jacoby, Robert W. Chestnut, Wayne D. Hoyer, David A. Sheluga, Michael J. Donahue, "Psychometric Characteristics of Behavioral Process Data: Preliminary Findings on Validity and Reliability," in J. Keith Hunt, ed., Advances in Consumer Research, Vol. 5, Association for Consumer Research, 1978.

Kelvin J. Lancaster, "A New Approach to Consumer Theory," Journal of Political Economy, Vol. 74, April, 1966, pp. 132-157.

Donald R. Lehmann, "Determinate of Choice Approach: A Research Paradigm," working paper, Columbia University, Graduate School of Business, 1978.

James H. Myers and Mark I. Alpert, "Determinant Buying Attitudes: Meaning and Measurement," Journal of Marketing, Vol. 32, October, 1968, pp. 13-20.

James H. Myers, "A Comparison of Two Methods for Determining Optimum Levels of Product Characteristics," in Advances in Consumer Research, Vol. 6, Urbana, Ill., Association for Consumer Research, 1978.

Jerry C. Olson and Aydin Mudderrisoglu, "The Stability of Responses Obtained by Free Elicitation: Implications for Measuring Salience and Memory Structure," in Advances in Consumer Research, Vol. 6, Urbana, Ill., Association for Consumer Research, 1978.

John W. Payne and E.K. Easton Ragsdale, "Verbal Protocols and Direct Observation of Supermarket Shopping Behavior: Some Findings and a Discussion of Methods," in J. Keith Hunt, ed., Advances in Consumer Research, Vol. 5, Association for Consumer Research, 1978.

Dov Pekelman and Subrata K. Sen, "Mathematical Programming Models for the Determination of Attribute of Attribute Weights," Management Science, Vol. 20, April, 1974, pp. 217-29.

John A. Quelch, "Behavioral Measurement of the Relative Importance of Attribute-Related Information Cues: The Case of Cold Breakfast Cereals," in Advances in Consumer Research, Vol. 6, Urbana, Ill., Association for Consumer Research, 1978.

M.J. Rosenberg, "Cognitive Structure and Attitudinal Affect," Journal of Abnormal and Social Psychology, Vol. 53, November 1956, pp. 376-372.

Edward J. Russo, "Eye Fixations Can Save the World: A Critical Evaluation and a Comparison Between Eye Fixations and Other Information Processing Methodologies," in J. Keith Hunt, ed., Advances in Consumer Research, Vol. 5, Association for Consumer Research, 1978.

V. Srinivasan and A.D. Shocker, "Linear Programming Techniques for Multidimensional Analysis of Preferences," Psychometrika, Vol. 38, No. 3, September, 1973, pp. 337-369.

William L. Wilkie, and Edgar A. Pessemier, "Issues in Marketing's Use of Multi-Attribute Attitude Models," Journal of Marketing Research, Vol. 10, November, 1973, pp. 428-441.

Peter Wright, "Consumer Choice Strategies: Simplifying Vs. Optimizing," Journal of Marketing Research, Vol. 12, February, 1975, pp. 60-67.

ON THE USE OF FORMULAS OF THE PREDICTIVE VALIDITY  
OF REGRESSION IN CONSUMER RESEARCH

Philippe Cattin, University of Connecticut

Abstract

A frequent measure of the predictive validity of a regression model is the crossvalidated correlation. Estimators of the population crossvalidated correlation can be used. A few such estimators can be found in the psychology literature. They are reviewed. The advantage of these estimators (over a sample crossvalidated correlation) is that they produce more precise estimates. An example of the use of these estimators in consumer research is presented.

Introduction

In the social sciences in general (and in consumer research in particular) it is often valuable to measure the predictive validity of a regression model. In most instances, one is interested in predicting the Y-value of an object compared to other objects (e.g. a consumer's utility for a product or for a concept) rather than the absolute Y-value of an object. Hence, relative prediction is what matters (rather than absolute prediction). An appropriate measure of predictive validity is the crossvalidated correlation (rather than the mean squared error of prediction). It can be estimated by splitting the available observations into an estimation sample and a validation (or holdout) sample (and computing the Pearson correlation between the actual Y-values of the objects in the validation sample with the Y-values predicted with the regression parameters estimated in the estimation sample). The resulting measure is a sample crossvalidated correlation (e.g. Goldberg, 1971; Scott and Wright, 1976). However, there are also estimators of the population crossvalidated correlation. These estimators are not well known. The purpose of this paper is to review them, to show their advantage over a sample crossvalidated correlation and to illustrate their use in consumer research.

Estimators Of The Population Cross-validated Correlation

Let  $Y_i$  be an observation on a criterion variable,  $X_{ij}$  an observation on one of  $p$  predictor variables ( $j = 1, \dots, p$ ),  $\alpha$  and  $\beta_j$  population parameters and let:

$$Y_i = \alpha + \sum_{j=1}^p X_{ij}\beta_j + \epsilon_i \quad (1)$$

be a regression model where  $\epsilon_i$  is the disturbance associated to observation  $i$ . The population parameters are usually unknown and can be estimated with  $N$  observations. If  $E(\epsilon_i) = 0$  ( $i=1, \dots, N$ ), and if  $E(\epsilon_i^2) = \sigma^2$  and  $E(\epsilon_i\epsilon_j) = 0$  for  $i \neq j$  ( $i, j = 1, \dots, N$ ), the Ordinary Least Squares (OLS) estimator of (1) is the Best Linear Unbiased Estimator (BLUE).

There are three mean squared errors that must be distinguished (Darlington, 1968, p. 173). By the same token, there are three correlations: (a) the sample correlation, (b) the correlation produced in the population by the true population weights (which we shall call population correlation), and (c) the correlation produced

in the population by the (regression) estimated weights (which we shall call population crossvalidated correlation). The squared sample correlation is:

$$R^2 = 1 - \frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^N (Y_i - \bar{Y})^2}$$

where  $\bar{Y}$  is the mean of the observations on  $Y$  and  $\hat{Y}_i$  is the regression estimate of observation  $i$ . The most common estimator of the squared population correlation is attributed to Wherry (1931):

$$\hat{\rho}^2 = 1 - \frac{N-1}{N-p-1} (1 - R^2). \quad (2)$$

This is not an unbiased estimator. However, Montgomery and Morrison (1973) have shown analytically that the maximum bias of (2) is only about .1/N.

Several estimators of the population crossvalidated correlation have been proposed. Recently, Schmitt, Coyle and Rauschenberger (1977) did a Monte Carlo study to compare a couple of estimators. In their simulation design, they did vary the population correlation, the average multicollinearity, the number of predictor variables and the number of observations available for estimation. The levels of each of these variables were representative of studies in the social sciences. They assumed that both criterion and predictor variables are random and normally distributed. The two formulas compared by Schmitt et al. were:

$$\hat{\rho}_c^2 = 1 - \frac{N-1}{N-p-1} \frac{N+p+1}{N} (1 - R^2) \quad (3)$$

and

$$\hat{\rho}_c^2 = 1 - \frac{N-1}{N-p-1} \frac{N-2}{N-p-2} \frac{N+1}{N} (1 - R^2). \quad (4)$$

These formulas were derived from two unbiased estimators of the population mean squared error of prediction, one assuming fixed predictor variables, the other random predictor variables (formulas (13) and (14) respectively in (Darlington, 1968, p. 173-174). However, as pointed out by Darlington, among others, an unbiased estimator of the population mean squared error of prediction cannot be translated into an unbiased estimator of the population crossvalidated correlation. Hence, (3) and (4) are not unbiased. The results obtained by Schmitt et al. indicate that both (3) and (4) underestimate the true population crossvalidated correlation. Moreover, even though the estimations carried by Schmitt et al. in their simulation assume random predictor variables, (3) seems to produce less biased results than (4) (and (4), not (3), is the formula that is derived from a mean squared error of prediction estimator that assumes random predictor variables). In fact, the average difference between actual and estimated squared population crossvalidated correlations (across all simulation results) is +.0080 with (3) while it is +.0176 with (4)

There are at least three other formulas that were carefully derived by Browne (1975), Burket (1964) and Srinivasan (1977). The derivation of Browne's formula assumes random predictor variables while the derivation of Burket's and of Srinivasan's formulas assumes fixed predictor variables. The results given by Schmitt et al. in their article are sufficient to compute estimates of the bias of these formulas. The results (Cattin, 1978) show that the average difference between actual and estimated squared population crossvalidated correlations (across all simulation results) is +.0029 with Browne's formula, +.0015 with Burket's and -.0018 with Srinivasan's (even though Browne's formula, like the estimations carried by Schmitt et al., is the only formula that assumes random predictor variables). These values are substantially closer to zero than the +.0080 obtained with (3). Browne's, Burket's and Srinivasan's formulas thus seem to be less biased.

Browne's formula is:

$$\hat{\rho}_c^2 = \frac{(N - p - 3)\hat{\rho}^4 + \hat{\rho}^2}{(N - 2p - 2)\hat{\rho}^2 + p} \quad (5)$$

where  $\hat{\rho}^2$  is the maximum of zero and (2) and  $\hat{\rho}^4$  is the  $(\hat{\rho}^2)^2 - [2p(1 - \hat{\rho}^2) / (N - 1)(N - p - 1)]$ . Browne has shown by Monte Carlo simulation that the bias of his estimator is relatively small even with a small N/p ratio, except for low correlations. Burket's formula is  $(\hat{\rho}^2)^2 / R^2$  where  $\hat{\rho}^2$  is an estimator of the squared population correlation. Replacing  $\hat{\rho}^2$  by (2) gives:

$$\hat{\rho}_c^2 = \frac{[(N - 1)R^2 - p]^2}{R^2(N - p - 1)^2} \quad (6)$$

(The estimate obtained with this formula should be set equal to zero when p is greater than  $(N-1)R^2$ ).

Srinivasan (1977, p. 64-65) recently argued that, if the mean squared error of prediction estimator that assumes fixed predictor variables (formula (13) in Darlington (1968, p. 173) is to be used to define a squared population crossvalidated correlation formula, two degrees of freedom must be subtracted (from  $N + p + 1$ ). This is because (a) the value of the intercept can be changed and (b) the slopes can be multiplied by any scalar without changing the resulting correlation between the criterion variable and the predictor variables. The resulting squared population crossvalidated correlation formula is:

$$\hat{\rho}_c^2 = 1 - \frac{N - 1}{N - p - 1} \frac{N + p - 1}{N} (1 - R^2) \quad (7)$$

This formula can be rationalized further with the following argument. If there is only one predictor variable (i.e.  $p=1$ ), a regression need not be run to get the sample correlation between the two (criterion and predictor) variables. A well-known formula can be used (e.g. formula 10.1.3 in (Winkler and Hays, 1975, p. 645)). Moreover, the population crossvalidated correlation and the population correlation are equal and can be estimated with (2) (where  $p=1$ ). Hence, (2) and (7) estimate the same thing when  $p=1$ , and since (7) actually reduces to (2), formula (7) makes sense. Furthermore, Srinivasan (1977, p. 67-69) has shown that the bias of his formula is relatively small by comparing the values obtained with his formula to those obtained by Schmidt (1970) by simulation. (The estimations carried by Schmidt assumed random predictor variables).

We have reported that (5), (6) and (7) are (slightly biased) estimators of the squared population crossvalidated correlation of a regression model (but seemingly less biased than (3) and (4)). The advantage of these formulas over a sample crossvalidated correlation is that they do not require that the available observations be split into two samples (estimation and validation). The resulting measure of predictive validity is more precise (even though it is slightly biased). This has been shown by simulation by Schmitt et al. (1977, p. 756-757). Moreover, this makes sense intuitively since (5), (6) or (7) takes all the available information into account at once, while a sample crossvalidated correlation cannot.

#### Choosing Among Multiattribute Models - An Illustration Of The Use Of (5), (6) or (7)

In regression one has to choose the form of the relationship between the criterion variable and any (interval or ratio scaled) predictor variable (e.g. Should it be linear, nonlinear? Should dummy variables be used?) Although there may be a priori reasons for selecting a function, one is often uncertain that the most appropriate function is used. If prediction is what matters, (5), (6) or (7) can be used to find out which of two (or more) potential functions seems to have more predictive validity.

The number of regression parameters corresponding to any predictor variable depends upon the assumed relationship with the criterion variable. If a linear function is assumed there is only one parameter. If a nonlinear function is assumed, there may be one or two (or even more) parameters. If dummy variables are used, the number of parameters is  $(k-1)$  where k is the number of levels the predictor variable takes; hence, it can be one, two or more. The sample correlation typically increases with the number of parameters to estimate. However, formula (5), (6) or (7) shows that the shrinkage between sample correlation and crossvalidated correlation increases with the number of parameters. Hence, the predictive validity of a model may or may not increase when an assumed linear function is replaced by (say) a quadratic function or by dummy variables.

An example will now be used to illustrate the use of formulas (5), (6) and (7). The data were taken out of an article by Green (1973). The predictor variables are the research, the teaching and the institutional contribution of a University Assistant Professor. Teaching and institutional contribution take on three levels: "below average", "average" and "superior". Research takes on the same three levels and "outstanding". This defines  $(3 \times 3 \times 4)$  36 hypothetical Assistant Professors. Green (1973, Table 1, p. 411) reports the response ratings of a subject in terms of his subjective probability (ranging from 0 to 100%) of recommending each Professor for a tenured faculty position. In a multiattribute context, these response ratings represent the observations on the criterion variable. A number of multiattribute models can be hypothesized depending upon the attribute utility function (including dummy variables) assumed for research, teaching and institutional contribution. Each model can in turn be estimated by regression.

For illustrative purposes let us consider two models: one using dummy variables for each attribute, the other assuming a linear function for each attribute. The dummy variables model has seven parameters, since  $\sum_i (k_i - 1) = 7$  (where  $k_i$  is the number of levels of attribute i).

On the other hand, the linear model has three parameters (one per attribute). Since outstanding, superior, average and below average correspond to the 98, 80, 50 and 20 percentile level respectively (as compared to all academics throughout the U.S. at similar career points in similar areas of specialization), we shall use these values as our observations on the predictor variables to estimate the linear model.

In a first step, each model was estimated by regression using all 36 observations. The squared sample correlation of the dummy variables model is slightly superior: .922 vs .918 (see Table 1A). However, the estimate of the squared population crossvalidated correlation of the linear model is somewhat higher than the corresponding estimate of the dummy variables model, whether we use (5), (6) or (7) (see Table 1A). Hence, the linear model seems to have more predictive validity. In other words, if we had another set of observations (provided by the judge who produced the 36 observations we used) we are likely to predict their Y-value more accurately with the linear model than with the dummy variables model.

In a second step, the 36 observations were split randomly into two subsamples of 18 observations each. Each subsample was used alternatively as estimation sample and as validation sample. The validation sample was used to compute a sample crossvalidated correlation. Moreover, (5), (6) and (7) were used to get estimates of the population crossvalidated correlation. The results are shown in Table 1B. When the first subsample is the estimation sample, the sample crossvalidated correlation and the estimate obtained with (5), (6) and (7) give an edge to the linear model. When the second subsample is the estimation sample, only the sample crossvalidated correlation and the estimate obtained with (5) give an edge to the linear model. However,

the average of the two gives an edge to the linear model whichever criterion is used. But moreover, the results obtained in the first step are more precise. All the available information is taken into account at once which leads to more precise estimates (as shown by simulation by Schmitt et al. (1977)).

The results also show that the estimates obtained with (5), (6) and (7) are quite close except in the case of the dummy variables model when only 18 observations are used for estimation (Table 1B). (When the second subsample is the estimation sample, the estimate obtained with (5) is .839 while it is .863 and .869 with (6) and (7) respectively). In this case the number of parameters is 8 (including the intercept). Hence, the ratio  $N/(n+1)$  is only 2.25. The estimates obtained with (5), (6) and (7) can differ substantially when this ratio is small.

#### Summary

In consumer research it is often valuable to know the predictive validity of a regression model. An appropriate measure is the crossvalidated correlation. Estimators of the population crossvalidated correlation can be used. A few such estimators can be found in the psychology literature. They were reviewed. The advantage of these estimators (over a sample crossvalidated correlation) is that they produce more precise estimates. An example of the use of these estimators in consumer research was presented.

#### References

M.W. Browne, "Predictive Validity of a Linear Regression Equation," British Journal of Mathematical and Statistical Psychology, 28(1975), 79-87.

TABLE 1

#### R-SQUARES AND SQUARED CROSSVALIDATED CORRELATIONS

	Dummy Variables Model (7 parameters)	Linear Model (3 parameters)
A. The estimation was carried on all 36 observations		
$R^2$	.922	.918
Browne's estimate	.881	.905
Burket's estimate	.883	.903
Srinivasan's estimate	.886	.906
B. The estimation was carried on one of two subsamples of 19 observations		
1. Subsample 1 was the estimation sample		
$R^2$	.942	.936
Browne's estimate	.839	.911
Burket's estimate	.863	.909
Srinivasan's estimate	.869	.914
Sample squared crossvalidated correlation	.862	.889
2. Subsample 2 is the estimation sample		
$R^2$	.965	.933
Browne's estimate	.903	.906
Burket's estimator	.917	.904
Srinivasan's estimator	.921	.909
Sample squared crossvalidated correlation	.895	.925

G.R. Burket, "A Study of Reduced Rank Models for Multiple Prediction," Psychometric Monographs, (1964, No. 12).

P. Cattin, "On Formulas of Crossvalidated Multiple Correlation," Working Paper: Center for Research and Management Development, University of Connecticut, (1978).

R.B. Darlington, "Multiple Regression in Psychological Research and Practice," Psychological Bulletin, 69(1968), 161-182.

L.R. Goldberg, "Five Models of Clinical Judgment: An Empirical Comparison Between Linear and Nonlinear Representations of the Human Inference Process," Organizational Behavior and Human Performance, 6(1971), 458-479.

P.E. Green, "On the Analysis of Interactions in Marketing Research Data," Journal of Marketing Research, 10(1973), 410-420.

D.B. Montgomery and D.C. Morrison, "A Note on Adjusting  $R^2$ ," Journal of Finance, 28(1973), 1009-1013.

F.L. Schmidt, "The Relative Efficiency of Regression and Simple Unit Predictor Weights in Applied Differential Psychology," Unpublished Doctoral Dissertation. West Lafayette, Indiana: Purdue University, 1970.

N. Schmitt, B.W. Coyle, and J. Rauschenberger, "A Monte Carlo Evaluation of Three Formula Estimates of Cross-validated Multiple Correlation," Psychological Bulletin, 84(1977), 751-758.

J.E. Scott, and P. Wright, "Modeling an Organizational Buyer's Product Evaluation Strategy: Validity and Procedural Considerations," Journal of Marketing Research, 13(1976), 211-224.

V. Srinivasan, "A Theoretical Comparison of the Predictive Power of the Multiple Regression and Equal Weighting Procedures," Research Paper No. 347: Graduate School of Business, Stanford University, (1977).

R.J. Wherry, Sr., "A New Formula for Predicting the Shrinkage of the Coefficient of Multiple Correlation," Annals of Mathematical Statistics, 2(1931), 440-457.

CONSTRUCT VALIDATION IN MARKETING: A COMPARISON OF METHODS IN ASSESSING THE VALIDITY OF THE AFFECTIVE, CONATIVE, AND COGNITIVE COMPONENTS OF ATTITUDES

George John (student), Northwestern University  
Torger Reve (student), Northwestern University

Abstract

The assessment of the validity of constructs in marketing are examined. An empirical comparison of different methods of establishing convergent and discriminant validity within the framework of multitrait-multimethod data analysis is made in the case of the affective, conative and cognitive components of attitude. Ostrom's (1969) data regarding attitudes toward the church was examined and revealed a lack of convergent and discriminant validity of the three components. The relative advantages and disadvantages of the different methods of analysis are compared and contrasted.

Dimensions of Construct Validity

One of the most basic issues in empirical, behavioral research in marketing, as well as in the other social sciences, is the notion of construct validity, i.e., "the extent to which an operationalization measures the concept which it purports to measure" (Zaltman et al., 1973, p. 44). In spite of the importance of construct validity from both a scientific and an applied point of view, construct validation has been largely ignored by most researchers in marketing and consumer research. Most studies in marketing tend to be single-measure, one-shot affairs with little or no attention to validity. Cook and Campbell (1976), in their extensive discussion of the various threats to validity confronting experimental and quasi-experimental research in field settings, equate construct validity with what experimental psychologists mean by confounding, i.e., "the possibility that the operational definition of a cause or effect can be construed in terms of more than one construct, all of which are stated at the same level of reduction" (1976, p. 238). Confounding in this context means that "what one investigator might interpret as a causal relationship between A and B, another investigator might interpret as a causal relationship between A and Y or between X and B or even between X and Y, and later experiments might support one or the other of these reinterpretations" (ibid., p. 238).

This leads to a consideration of construct validity in terms of three empirical dimensions (Campbell and Fiske, 1959; Campbell, 1960; Kerlinger, 1973; Cook and Campbell, 1976, Bagozzi, 1977): (1) Convergent validity, i.e., the degree to which two or more attempts to measure the same concept through maximally different methods are in agreement, that is, intermethod convergence. (2) Discriminant validity, i.e., the degree to which a concept differs from other concepts. (3) Nomological validity, i.e., the degree to which predictions from formal theoretical networks containing the concept in question are confirmed. Bagozzi (1978) specifies two additional criteria which need to be met in order to demonstrate construct validity: (4) Semantic criteria, i.e., theoretical and observational meaningfulness. (5) Operational empirical criteria, i.e., internal consistency and reliability. The theoretical meaningfulness of a construct refers to the nature and internal consistency of the language used to represent the construct, while operational meaningfulness refers to the conceptual relationships between a theoretical variable and its operationalizations(s).

Comparing this full set of five criteria for achieving construct validity with the type of criteria usually found in published marketing studies clearly shows the current insufficiencies. At best, three types of criteria are applied in evaluating the measures used in marketing studies (c.f., Heeler and Ray, 1972): (1) Face or consensus validity, i.e., the degree to which a measure "looks as if" it should indicate a particular variable or concept. An example would be the prevalent use of recall measures as an indicator of advertising exposure and reception. (2) Predictive or concurrent validity, i.e., the extent to which a particular measure predicts other criterion measures (Zaltman et al., 1973, p. 44). An example would be the degree of correlation between post ad recall and attitude change. Another interpretation of the criterion is that the measure has been used extensively by others, which thus may lead to further strengthening of the "conventional wisdom." (3) Reliability, i.e., the degree of agreement between two efforts to measure the same concept through maximally similar methods (Campbell and Fiske, 1959), that is, intramethod convergence. Testing for reliability shows the extent to which scores on specific measures may be caused by irrelevant factors.

Most marketing studies seem to be lacking in terms of demonstrating either convergent or discriminant validity of the concepts used. Even when such validation is attempted, the usual practice is to use unstructured or ad hoc procedures, which are generally factor analytic or correlational in nature. A discussion of some of these factor analytic approaches follows in the next section.

Ad Hoc Procedures for Assessing Construct Validity

The various empirical procedures for assessing construct validity which will be dealt with below, are basically confined to an analysis of convergent and discriminant validity.

Factor Analysis for Assessing Convergent Validity

The context for using a simple factor analytic method for assessing convergent validity is where a construct has been operationalized in terms of multiple measures, e.g., when several items in a questionnaire are taken to represent the same variable. The obtained measures for each variable can be factor analyzed to determine whether all items merit inclusion in the variable. Loosely, the criterion for excluding a particular item is that it exhibits little loading on the first factor extracted. If on the other hand, all measures load heavily on the first factor extracted, one can usually conclude that the items are all dimensions of the same underlying factor, that is, the variable that one intended to measure. An extension of this approach is to use the obtained factor scores for each subject in order to weight the items empirically when computing the composite score for the variable.

Factor Analysis for Assessing Discriminant Validity

The context for using a simple factor analytic method for assessing discriminant validity is where one is to show that two or more theoretical constructs are



are distinct and separate. A case in question would be where several constructs are included in a study, and each construct is operationalized by multiple measures. Evidence for discriminant validity is obtained if the measures for separate constructs do not load on the same factors. An example from marketing of this method for establishing discriminant validity is a study of distribution channels by Lusch (1976). Trying to distinguish between coercive and noncoercive sources of power, he included 16 sources which were thought to be noncoercive and 6 sources which were thought to be coercive and subjected his data obtained to common factor analysis. Of the five factors extracted, the first was the coercive factor, while the other four factors were noncoercive. In no cases did coercive and noncoercive sources load on the same factor, which was then taken as evidence for discriminant validity of the two constructs. A possible confounding in this particular case, however, was that different types of questions and different types of scales were used for coercive and noncoercive sources of power. Although the differences in scales are probably partly resolved through normalization of the data, the differences in types of questions may partially account for the loading on separate factors, that is, the extraction of a method factor rather than actual discriminant validity of the two constructs.

#### Factor Analysis for Assessing Both Convergent and Discriminant Validity

The two previous factor analytic procedures could easily be combined in the case where both multiple constructs and multiple measures of each construct are utilized. In terms of the example discussed above (Lusch, 1976), a simultaneous demonstration of discriminant and convergent validity could have been made if the five factors extracted corresponded to one coercive power factor, and to four distinct subcategories of noncoercive power (French and Raven, 1959). The above methods for assessing convergent and discriminant validity are ad hoc approaches in that there is no structured procedure for disentangling the method and trait variance. A solution to such problems may be found when turning to multitrait-multimethod procedures for assessing construct validity as first suggested by Campbell and Fiske (1959). Some recent approaches for assessing construct validity using multitrait-multimethod data will then be presented and applied to the attitude construct.

#### Multitrait-Multimethod Procedures for Assessing Construct Validity

Campbell and Fiske (1959), introduced the multitrait-multimethod matrix (MTMM-matrix) as a tool for assessing convergent and discriminant validity. The first reported attempt to utilize the MTMM methodology in marketing were made by Myers (1968) and Ray (1968). In spite of the large number of present and potential applications of MTMM analysis in marketing suggested by Heeler and Ray (1972), this methodology for assessing construct validity has not yet been widely used within marketing and consumer research. Some of the possible extensions of the MTMM approach to construct validation as suggested by several marketing studies include multirespondent-multimethod matrices (Davis, 1971), multitrait-multicontext matrices (Robertson and Myers, 1969; Myers and Robertson, 1972), and even multiitem-multi-product matrices (Sears, 1969). Davis (1971) in his study of family decision making measured purchase influence of husbands and wives using four different measures, while Myers and Robertson (1972) developed the data for opinion leadership traits across a large number of contexts, but without putting the data into a full matrix format. Several methods for analyzing multitrait-multimethod data exist (Schmitt et al., 1977).

In this paper three different methods will be outlined: (1) Campbell and Fiske counting procedure, (2) Jöreskog Restricted Maximum Likelihood Factor Analysis, (3) Jackson's Two Stage Principal Component Analysis.

#### Campbell and Fiske Analysis

The multitrait-multimethod approach tries to determine convergent and discriminant validity through an analysis of the pattern of correlations among two or more traits as measured by two or more methods. An example of a multitrait-multimethod matrix where three traits (A,B,C) are measured by three methods (1,2,3) is shown in Figure 1. The MTMM matrix contains the intercorrelations among the observed measures which comprise the available observed data. Let  $r_{(jk)(j'k')}$  represent the correlation between the measure  $jk$  and the measure  $j'k'$  where  $j$  designates trait and  $k$  designates method. The MTMM matrix can then be described by the following partitions: 1. The monotrait-monomethod values are composed of the elements  $r_{(jk)(j'k')}$  where  $j=j'$  and  $k=k'$ .

These main diagonal elements of the MTMM matrix represent the reliabilities of the measures, either in terms of stability or internal consistency. If these correlations are found to be sufficiently high, further validation can be undertaken. 2. The heterotrait-heteromethod triangles are composed of the elements  $r_{(jk)(j'k')}$  where  $j \neq j'$  and  $k \neq k'$ . 3. The validity diagonals or monotrait-heteromethod values are composed of the elements  $r_{(jk)(j'k')}$  where  $j = j'$  and  $k \neq k'$ . 4. The

heterotrait-monomethod triangles are composed of the elements  $r_{(jk)(j'k')}$  where  $j \neq j'$  and  $k = k'$ . To assess convergent and discriminant validity in terms of the pattern of correlations as indicated by the MTMM matrix as described above, Campbell and Fiske (1959, pp. 82-83) suggested four types of comparisons of correlation coefficients, sometimes referred to as the counting method: 1. Evidence of convergent validity is established by examining the correlations between different measures of the same traits. This means that the elements in the validity diagonal value should be higher than the values lying in its column and row in the heterotrait-heteromethod triangles. For a given variable, its validity diagonal value should be higher than its corresponding value in the heterotrait-monomethod triangles. The same pattern of trait interrelationships should be shown in all of the heterotrait triangles of both monomethod and heteromethod blocks. As can be shown from the simple case of three traits and three methods, a full utilization of the Campbell and Fiske criteria will lead to a large number of comparisons of correlations in order to establish both convergent and discriminant validity. Furthermore the criteria become ambiguous when only some of the comparisons fulfill the criteria while others do not. A simple counting procedure of the number of comparisons that satisfy each criterion to the number of comparisons that do not (e.g., Kothanadapani, 1971) may lead to false conclusions as to convergent and discriminant validity (Bagozzi, 1978).

The Campbell and Fiske counting criteria further assume that (1) the traits and methods are uncorrelated, (2) the methods are only minimally intercorrelated (Campbell and Fiske, 1959; Alwin, 1974; Kalleberg and Kluegel, 1975). In practice, however, this is rarely true (c.f., Jackson, 1969; Alwin, 1974). Consider for instance, the violation of assumption (2). One would obtain spuriously high correlations in the monotrait-heteromethod cases, thus, leading to a misleading conclusion of high convergent validity. Consequently, various analytic schemes have been proposed as alternatives to the original Campbell and Fiske criteria.

One class of models is the restricted maximum-likelihood factor analysis models. Boruch and Wolins (1970) propose the model

$$Y_{ijk} = B_{jk} X_{ij} + C_{jk} X_{ik} + \epsilon_{ijk} \quad (1)$$

where  $Y_{ijk}$  = observation on  $i^{\text{th}}$  subject on hypothetical factor associated with trait  $j$  (2)

$X_{ik}$  = score of  $i^{\text{th}}$  subject on hypothetical factor associated with method  $k$  (3)

$\epsilon_{ijk}$  = error associated with  $i^{\text{th}}$  subject, trait,  $j$ , method  $k$  (4)

The covariance matrix of the observations can be written as  $\Sigma = \Lambda \Phi \Lambda' + \Psi$  where  $\Phi$  is the factor inter-correlation matrix and  $\Psi$  is the covariance matrix of the error terms. If the data is to exhibit evidence of convergent and discriminant validity then it should be possible to fit such a model to the data which is also a unique solution in that it cannot be rotated as usual factor analysis results can be treated. We obtain such a unique solution by specifying certain restrictions on the parameters of the various matrices. Consider the loading matrix in Figure 2.

FIGURE 1

MULTITRAIT-MULTIMETHOD MATRIX OF CORRELATIONS FROM OSTROM (1969)

	Thurstone			Likert			Guttman		
	A	B	C	A	B	C	A	B	C
T	A	-							
	B	.57	-						
	C	.63	.62	-					
L	A	.71	.59	.68	-				
	B	.67	.67	.71	.79	-			
	C	.69	.62	.72	.79	.81	-		
G	A	.54	.39	.49	.58	.51	.56	-	
	B	.59	.61	.58	.60	.69	.60	.43	-
	C	.63	.50	.63	.69	.67	.71	.49	.56

The decimal point of each correlation has been omitted. N = 189, A = affective, B = behavioral, c = cognitive.

Each column represents a trait or a method factor and the rows represent the different trait-method measures. Now if a particular trait is present in a measure, then the loading of that measure on its trait factor is left unrestricted. The same logic is applied to the methods in each measure. However, all other loadings are restricted to be zero. This is the model which would hold if the traits did indeed possess discriminant and convergent validity. The  $\Phi$  matrix is left unrestricted. However, we may restrict the parameters of  $\Phi$  such that the trait and method factors are uncorrelated if we feel that such an assumption is warranted in a particular case.

A unique solution (or identifiability) may not always be possible and the necessary and sufficient conditions for such a solution existing have not been worked out for the general case of this class of models. As for

significance tests, a  $X^2$  statistic based on a likelihood ratio is available which compares the estimated correlation matrix produced by the model with the actual correlation matrix. This statistic can also be used to test particular parameters in a model by estimating one model as a restricted subset of another model (Jöreskog, 1971). In essence, then, in this approach, we are hypothesizing a particular factor structure and testing it. Jöreskog (1967, 1971) has developed an efficient estimation procedure using maximum-likelihood

procedures for this kind of factor analytic model (i.e., the RMLFA model).

The analysis strategy, then, is to attempt to search for a reasonable model of the sort described which fits the data to an acceptable degree thus enabling us to conclude that discriminant validity and convergent validity have been achieved.

Jackson's Two Stage Principal Components Method

Jackson (1969) proposed an alternate approach to the analysis of multimethod multitrait data. The previous model ( $\Sigma = \Lambda \Phi \Lambda' + \Psi$ ) assumes that the data is comprised of a set of trait and method factors. If this is indeed the case, a principal axes analysis of the correlation should yield these trait and method factors. However, a straightforward application of a principal components analysis would yield mixed method and trait factors because of the trait-method factor intercorrelations. In order to derive estimates of the variance accounted for by the trait and method factors, Jackson (1975) proposed the following procedure. Essentially, the procedure is a two-stage factor analysis where the first stage consists of an orthogonalization of the mono-method heterotrait submatrices via a full principal components analysis. Then, a cross-products matrix of

FIGURE 2

A MATRIX

		Beh. Int			Cognition	Thurstone	Likert	Guttman
		A	B	C				
Thurstone	A1	$\lambda_{11}$	0	0	$\lambda_{14}$	0	0	
	B1	0	$\lambda_{22}$	0	$\lambda_{24}$	0	0	
	C1	0	0	$\lambda_{33}$	$\lambda_{34}$	0	0	
Likert	A2	$\lambda_{41}$	0	0	0	$\lambda_{45}$	0	
	B2	0	$\lambda_{52}$	0	0	$\lambda_{55}$	0	
	C2	0	0	$\lambda_{63}$	0	$\lambda_{65}$	0	
Guttman	A3	$\lambda_{71}$	0	0	0	0	$\lambda_{76}$	
	B3	0	$\lambda_{82}$	0	0	0	$\lambda_{86}$	
	C3	0	0	$\lambda_{93}$	0	0	$\lambda_{96}$	

factor scores is constructed and subjected to a further principal components analysis. In this second stage, the factor loading matrix that is obtained should exhibit "clean" trait factors if convergent and discriminant validity exists. For the technical details of the method see Jackson (1975). The justification of this procedure follows from the fact that the restriction of orthogonality we imposed via a principal components analysis on the individual traits within a method from being reflected in the constructed matrix of component scores. Thus, if we define method variance as variance unique to a particular method of measurement, then such method variance will not be present in common factors derived from the intercorrelation matrix of factor scores. In a similar fashion, trait variance can be eliminated and method factors can be extracted by rearranging the measures so that all methods of measuring each trait are grouped together in the diagonal submatrices of the MTMM matrix.

The Tripartite Classification of the Attitude Construct

The MTMM matrix analyzed in this study attempts to establish the convergent and discriminant validity of the tripartite classification of the attitude construct.

The attitude construct itself has been the center of considerable controversy especially in its relation to behavior. At one extreme, it has been argued (eg., Doob, 1947) that attitudes are merely overt behavioral

responses and that its existence separate from behavior cannot be justified. On the other hand, the evidence from attitude-behavior studies seems to indicate a lack of consistent, strong relationships between attitude and behavior (Wicker, 1969). Generally, though, it is held (Fishbein and Ajzen, 1975; Greenwald, 1968) that attitudes are causally linked to behavior. If we are to attempt to test this linkage, it is imperative to define the attitude construct and measure its components. Otherwise, our conclusions regarding attitude-behavior consistency will remain equivocal as it becomes difficult, if not impossible, to pin down the locus of the cause of the observed relationship (or lack thereof). There are two broad classes of models of the attitude construct. The first of these is the expectancy value model (eg., Fishbein and Ajzen, 1975) which holds that an attitude toward an object or act consists of two components, viz. cognitions of values of outcomes or goals associated with the object and the beliefs that the object will help or hinder the attainment of these outcomes or goals. Then, these cognitions and beliefs are combined in some fashion (usually a linear additive model is used) to determine one's attitude.

The major competing model is the tripartite classification scheme (Rosenberg, 1960; Insko and Scholpler, 1967). Here, three distinguishable classes of concepts regarding an object are held to exist. These are the cognitive, affective and conative (or behavioral intention) components. Briefly, the cognitions are the statements of beliefs or facts regarding the object, while the affective component is the emotional content of thoughts regarding the object. The affective component is also regarded as being evaluative/directional with respect to the object. The conative component is the action tendency one has toward the object. No hierarchical ordering of these components is presumed about the linkages between these categories. Each component can affect behavior separately. More importantly different antecedent processes are associated with each of them (eg., Greenwald, 1968).

There have been two major attempts at validating the scheme by collecting data within an MTMM format (Ostrom, 1969; Kothandapani, 1971). Some evidence for convergent and discriminant validity of the components were found by both authors but the results must be regarded as rather tentative because of the limitations of the analysis used in the two studies. In this reanalysis, the data from Ostrom's study is investigated for convergent and discriminant validity.

#### Data Collection

Ostrom (1969) collected data about attitudes toward the church. From his sample of 189 undergraduate students, he constructed four measures of each attitude component. We shall examine the MTMM matrix formed from three of the methods (viz. a Thurstone scale, a Likert scale and a Guttman scale). A complete description of the actual scale items, the procedures used in scale construction and the collection of data is given in Ostrom (1969). Figure 1 shows the MTMM matrix.

#### Empirical Results

Campbell-Fiske Counting Method: Ostrom applied the counting criteria to his MTMM matrix and concluded that it did provide evidence of convergent and discriminant validity of the components. We did not duplicate his analysis but shall note that this result can be compared to those obtained from the other two methods.

Jöreskog's Restricted Maximum Likelihood Factor Analysis: The parameters of the various hypothesized models were estimated by the COFAMM computer program developed by Jöreskog and Sörbom (1976). The first model that

was estimated was a six factor model (three trait and three method factors). Figure 3 shows the estimated  $\Lambda$ ,  $\Phi$ , and  $\Psi$  matrices. The goodness-of-fit statistic shows an excellent fit ( $p = .98$ ). However, an examination of the parameters reveals certain problems. For instance, two factor loading parameters are greater than unity while one specific variance estimate is negative. These boundary problems indicate that the model is an unacceptable solution and can at best be used to give us clues towards modifying the model. Specifically, the high factor intercorrelations suggested that a more parsimonious solution may be possible. Table 1 summarizes the various models and their solutions. It should be noted that boundary problems were present for almost every single model although the  $X^2$  statistic indicated an acceptable fit in almost every case. The two models that were characterized by acceptable parameter estimates are shown in Figure 4 and Figure 5. Figure 4 displays a three trait two method model with restrictions in the  $\Lambda$  matrix such that the trait and method factors are uncorrelated. The degree of fit is acceptable ( $p = .72$ ), and one would assume that convergent and discriminant validity had been established. However, an examination of the trait factor intercorrelations reveals extremely high values. Specifically,  $\phi_{13} = .935$ ,  $\phi_{13} = .935$ ,  $\phi_{23} = .901$  and  $\phi_{12} = .857$ . This means that the trait factors are highly collinear and suggests that the traits (i.e., the attitude components) are virtually indistinguishable from each other. Consequently, two of the components (affect and cognition) were collapsed into a single factor and the resulting model was estimated. Figure 5 shows that the fit is not unacceptable ( $p = .58$ ) and that this model cannot be rejected as an adequate representation of the data.

This final model does not support a conclusion of three distinct, discriminable traits. Convergent validity, however, is present when considering the magnitudes of the loadings in the model estimates in Figure 4 and Figure 5. It should be noted that the  $X^2$  tests are not independent of each other, and, ideally, we would like to have a hold-out sample for confirmation of the final model.

FIGURE 3

MODEL 1 in TABLE 1

LAMBDA ( $\Lambda$ ) MATRIX	
	1 2 3 4 5 6
1 AFF-TH	.553 0.000 0.000 .366 0.000 0.000
2 BEH-TH	0.000 .065 0.000 .817 0.000 0.000
3 COG-TH	0.000 0.000 .486 .378 0.000 0.000
4 AFF-LIK	.410 0.000 0.000 0.000 .556 0.000
5 BEH-LIK	0.000 .150 0.000 0.000 .939 0.000
6 COG-LIK	0.000 0.000 .509 0.000 .423 0.000
7 AFF-GUT	.494 0.000 0.000 0.000 0.000 -.262
8 BEH-GUT	0.000 1.036 0.000 0.000 0.000 -.920
9 COG-GUT	0.000 0.000 1.274 0.000 0.000 .560

PHI ( $\Phi$ ) MATRIX	
	1 2 3 4 5 6
1	1.000
2	.033 1.000
3	.857 -.085 1.000
4	.595 .031 .789 1.000
5	.712 -.092 .891 .864 1.000
6	-.625 .414 -.808 -.719 -.790 1.000

PSI ( $\Psi$ ) MATRIX	
	1 2 3 4 5 6
1	.32 .325 .332 .199 .122 .179
7	.8 .9
	.525 -.131 .216

Factor 1:	Affect trait factor
Factor 2:	Behavioral Intention trait factor
Factor 3:	Cognition trait factor
Factor 4:	Thurstone method factor
Factor 5:	Likert method factor
Factor 6:	Guttman method factor

FIGURE 4

MODEL 6 IN TABLE 1  
 $\chi^2(14)=10.16$  (p=.75)

LAMBDA					
	1	2	3	4	5
1 AFF-TH	.760	0.000	0.000	.348	0.000
2 BEH-TH	0.000	.779	0.000	.276	0.000
3 COG-TH	0.000	0.000	.712	.403	0.000
4 AFF-LIK	.656	0.000	0.000	.624	0.000
5 BEH-LIK	0.000	.683	0.000	.646	0.000
6 COG-LIK	0.000	0.000	.664	.622	0.000
7 AFF-GUT	.534	0.000	0.000	0.000	.375
8 BEH-GUT	0.000	.737	0.000	0.000	.306
9 COG-GUT	0.000	0.000	.620	0.000	.509

PHI					
	1	2	3	4	5
1	1.000				
2	.857	1.000			
3	.935	.901	1.000		
4	0.000	0.000	0.000	1.000	
5	0.000	0.000	0.000	.936	1.000

PSI						
	1	2	3	4	5	6
1	.301	.394	.331	.181	.118	.174
7						
8						
9						
	.574	.366	.359			

Factor 1: Affect factor  
 Factor 2: Beh. Intention factor  
 Factor 3: Cognition factor  
 Factor 4: Thurstone + Likert factor  
 Factor 5: Guttman factor

FIGURE 5

MODEL 7 IN TABLE 1  
 $\chi^2(16)=14$  (p=.58)

LAMBDA				
	1	2	3	4
1 AFF-TH	.536	0.000	.657	0.000
2 BEH-TH	0.000	.379	.647	0.000
3 COG-TH	.270	0.000	.749	0.000
4 AFF-LIK	.304	0.000	.828	0.000
5 BEH-LIK	0.000	.268	.883	0.000
6 COG-LIK	.213	0.000	.882	0.000
7 AFF-GUT	.349	0.000	0.000	.546
8 BEH-GUT	0.000	.591	0.000	.614
9 COG-GUT	.301	0.000	0.000	.733

PHI				
	1	2	3	4
1	1.000			
2	.608	1.000		
3	0.000	0.000	1.000	
4	0.000	0.000	.981	1.000

PSI						
	1	2	3	4	5	6
1	.281	.438	.366	.222	.149	.177
7						
8						
9						
	.580	.275	.372			

Factor 1: Affect-Cognition factor  
 Factor 2: Beh. Intention factor  
 Factor 3: Thurstone + Likert factor  
 Factor 4: Guttman factor

Jackson's Two Stage Principal Components Analysis

The present authors applied the two stage technique to this data and the loading matrix (after Varimax rotation) of the second stage is shown in Figure 6. The three factor solution in Figure 6 was chosen on the basis of an eigenvalue cutoff rule (i.e., less than unity). Consider the first factor, A1, B3 and C2 all load heavily on this factor. Since these are measures of affect, conation and cognition respectively, it is clear that we do not have a clean 'trait' factor. In the second factor, B1, A2, A3 are the measures that load heavily on it. Again, the presence of the conation measure, B1, prevents us from labelling it as an 'affect' factor. C1, B2 and C3 contribute heavily to the third factor and again, we do not seem to have a clean trait factor.

The conclusion to be drawn from this analysis is that we cannot conclude that the traits or components are sufficiently different and distinguishable from each other to say that convergent and discriminant validity have been established.

TABLE 1

RESTRICTED MAXIMUM LIKELIHOOD FACTOR ANALYSIS MODELS

Model:	Restrictions	Result
1 3 traits and 3 method factors	$\Delta$ is restricted $\phi$ free $\psi$ free	$\chi^2(3) = 118$ (p = .98) 2 $\Delta$ parameters $> 1$ 1 $\psi$ parameter $< 0$ High trait correlations ( $> .90$ ) High method correlations ( $> .90$ )
2 3 traits and 3 method factors	$\Delta$ is restricted $\phi$ is restricted so that trait factors are uncorrelated with method factors $\psi$ free	failed to converge
3 1 trait and 3 method factors	$\Delta$ is restricted $\phi$ is restricted so that the trait factor is uncorrelated with method factors $\psi$ free	failed to converge
4 3 traits and 2 method factor (Thurstone and Likert methods were collapsed)	$\Delta$ is restricted $\phi$ free $\psi$ free	failed to converge
5 3 traits and 2 methods (Guttman was collapsed into the other two method factors)	$\Delta$ is restricted $\phi$ is restricted so that trait factors are uncorrelated with method factors $\psi$ free	$\chi^2(11) = 4.7$ (p = .94) 1 $\Delta$ parameter $> 1$ 1 $\psi$ parameter $< 0$
6 3 traits and 2 methods (Thurstone and Likert methods were collapsed)	$\Delta$ is restricted $\phi$ is restricted so that trait factors are uncorrelated with method factors $\psi$ is free	acceptable solution See Figure 6
7 2 traits and 2 methods (Affect and cognition were combined; Thurstone and Likert were combined)	$\Delta$ is restricted $\phi$ is restricted so that trait factors are uncorrelated with method factors $\psi$ free	acceptable solution See Figure 7.

FIGURE 6

LOADING MATRIX OF THE SOLUTION OF THE 2-STAGE PRINCIPAL COMPONENTS ANALYSIS. VARIMAX ROTATED.

	Factor 1	Factor 2	Factor 3
Thurstone A1	.73	.04	-.04
B1	.12	.76	.01
C1	.26	.12	.71
Likert A2	.09	.77	.16
B2	-.12	-.06	.59
C2	.77	.01	.24
Guttman A3	-.01	.67	.07
B3	.81	.15	.02
C3	.09	.24	.68

Discussion

The two factor analytic methods concur in that there is a lack of convergent and discriminant validity. The counting procedure, on the other hand, finds support for both types of validity. These apparent contradictions can be traced back to the assumptions of each analysis.

In the counting procedure, the assumptions that the trait and method factors are uncorrelated is probably untenable for this data. Figure 3 shows the estimated correlations of the factors in the matrix and it can be seen that they are very high. In fact, two of the method factors were collapsed into a single factor in the later models. These results mean that the conclusions drawn from the counting procedure may be spurious.

The restricted maximum likelihood approach is an elegant analytic scheme, but it suffers from a number of operational and computational drawbacks. As we indicated previously, a large number of the fitted models have parameter estimates that are clearly meaningless (eg., negative specific variances). These problems leave the status of these models unsettled. This is also true of those models which failed to converge to a solution. The result of these problems is that the final fitted model is the result of considerable iterative model fitting. This can result in different interpretations being drawn from the same data. Bagozzi (1978) analyzed this data and estimated a model that is virtually identical to Figure 6 and concluded that convergent and discriminant validity had been established. However, as we saw previously, we did not accept this model as the final model because of the unacceptably high intercorrelation between the trait factors.

The Jackson procedure, in contrast, is computationally more appealing and is subject to fewer problems of the type mentioned above. However, at the theoretical level, it suffers from some potentially serious drawbacks. For instance, in the first stage, we orthogonalize the diagonal submatrices to extract trait factors. However, we are not disentangling the variance contribution from the trait and method sources; i.e., there is a possibility of confounding these sources of variation. Further, the principal components model does not allow for specific or residual variance. Thus, if we are dealing with measures with relatively large amounts of specific variance, this could well obscure the convergence and discriminability of the traits.

Table 2 summarizes the relative advantages and disadvantages of the methods. The best strategy to follow in the face of these various drawbacks of these different techniques is to proceed to use them simultaneously. In this way, the credibility of the findings is enhanced if different techniques permit one to draw the same conclusion. In the present study, both factor analytic techniques arrive at the same conclusion that it is not possible to establish the validity of the tripartite classification scheme for this data. We should mention, however, that this was a very stringent test of the scheme as we were dealing with attitudes toward an object where we would expect a great deal of consistency and collinearity between the traits. This is because attitudes towards the church are formed over a long period of time and is therefore likely to exhibit a great deal of stability and consistency. In spite of this, we did establish the presence of two separate traits on the Jöreskog models; this indicates that it is possible that the three components will be discriminable in other less stringent situations.

TABLE 2

COMPARISON OF METHODS OF ANALYZING MIMM DATA

<u>Issue</u>	<u>Campbell-Fiske Counting Method</u>	<u>Restricted Max.Likelihood Factor Analysis</u>	<u>Jackson's 2 stage Procedure</u>
Disentangling of trait and method variance	No Yes (if trait and methods are uncorrelated)	Yes	Confounded in first stage
Estimates of variance from each source(ie. trait,method and error)	No	Yes	No
Specificity of evaluation of presence of conv. and disc. validity	Vague. (% of elements that meet criteria)to total)	Specific. $X^2$ test of goodness-of-fit,	Less specific. 'Clean' trait factors in second stage.
Location of individuals on traits(ie. factor scores)	No	Yes	Yes
Computational ease	Easy	Difficult. Requires iterative model fitting	Easy.

## References

- D. Alwin, "An Analytic Comparison of Four Approaches to the Interpretation of Relationships in the Multi-trait-Multimethod Matrix," in H. Costner (ed.), Sociological Methodology 1973-1974 (San Francisco: Jossey-Bass, 1974)
- R.P. Bagozzi, "The Construct Validity of the Affective, Behavioral and Cognitive Components of Attitude by Analysis of Covariance Structures," Multivariate Behavioral Research, 13(January, 1978), 9-31.
- R.F. Boruch and L. Wolins, "A Procedure for Estimation of Trait, Method, and Error Variance Attributable to a Measure," Educational and Psychological Measurement, 30(1970), 547-574.
- D.T. Campbell, "Recommendations for APA Test Standards Regarding Construct, Trait, or Discriminant Validity," American Psychologist, 15(1960), 546-553.
- D.T. Campbell and D.W. Fiske, "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix," Psychological Bulletin, 56(March, 1959), 81-105.
- T.D. Cook and D.T. Campbell, "The Design and Conduct of Quasi-Experiments and True Experiments in Field Settings," in M.D. Dunnette (ed.), Handbook of Industrial and Organizational Psychology (Chicago: Rand McNally, 1976), 223-326.
- H.J. Davis, "Measurement of Husband-Wife Influence in Consumer Purchase Decisions," Journal of Marketing Research, 8(August, 1971), 305-312.
- L.W. Doob, "The Behavior of Attitudes," Psychological Review, 54(1947), 135-156.
- M. Fishbein and I. Ajzen, Belief, Attitude, Intention and Behavior (Reading, Mass.: Addison-Wesley, 1975)
- J.R.D. French and B. Raven, "The Bases of Social Power," in D. Cartwright (ed.), Studies in Social Power (Ann Arbor, Institute for Social Research, University of Michigan, 1959), 150-167.
- A.G. Greenwald, "On Defining Attitude and Attitude Theory," in A.G. Greenwald, T.C. Brock and T.M. Ostrom (eds), Psychological Foundations of Attitudes (New York: Academic Press, 1968), 361-388.
- R.M. Heeler and M.L. Ray, "Measure Validation in Marketing," Journal of Marketing Research, 9(November, 1972), 361-370.
- R.M. Heeler, T.W. Whipple and T.P. Hustad, "Maximum Likelihood Factor Analysis of Attitude Data," Journal of Marketing Research, 14(February, 1977), 42-51.
- C.A. Insko and J. Schopler, "Triadic Consistency: A Statement of Affective-Cognitive Consistency," Psychological Review, 74(1967), 361-376.
- D.N. Jackson, "Multimethod Factor Analysis in the Evaluation of Convergent and Discriminant Validity," Psychological Bulletin, 72(July, 1969), 30-49.
- D.N. Jackson, "Multimethod Factor Analysis," Multivariate Behavioral Research, 10(July, 1975)
- K.G. Jöreskog, "Some Contributions to Maximum Likelihood Factor Analysis," Psychometrika, 32(December, 1967), 443-482.
- K.G. Jöreskog, "Statistical Analysis of Sets of Congeneric Tests," Psychometrika, 36(June, 1971), 109-133.
- K.G. Jöreskog and D. Sörbom, COFAMM: A Fortran IV Program (Chicago: National Resources, Inc., 1976)
- A.J. Kalleberg and J.R. Kluegel, "Analysis of the Multi-trait-Multi-Method Matrix: Some Limitations and an Alternative," Journal of Applied Psychology, 60(1975), 1-9.
- F.N. Kerlinger, Foundations of Behavioral Research, 2nd Edition (New York: Holt, Rinehart and Winston, 1973)
- V. Kothandapani, "Validation of Feeling, Belief, and Intention to Act as Three Components of Attitude and Their Contribution to Prediction of Contraceptive Behavior," Journal of Personality and Social Psychology, 19 (September, 1971), 321-333.
- R.F. Lusch, "Sources of Power: Their Impact on Intra-channel Conflict," Journal of Marketing Research, 13 (November, 1976), 382-390.
- J.G. Myers, Consumer Image and Attitude (Berkeley: Institute of Business and Economic Research, University of California at Berkeley, 1968)
- J.G. Myers and T.S. Robertson, "Dimensions of Opinion Leadership," Journal of Marketing Research, 9(1972)
- T.M. Ostrom, "The Relationship Between the Affective, Behavioral, and Cognitive Components of Attitudes," Journal of Experimental Social Psychology, 5(1969), 12-30.
- M.L. Ray, "Neglected Problems (opportunities) in Research: The Development of Multiple and Unobtrusive Measures," Fall Conference Proceedings (Chicago: American Marketing Association, 1968), 176-183.
- T.S. Robertson and J.G. Meyers, "Personality Correlates of Opinion Leadership and Innovative Buying Behavior," Journal of Marketing Research, 6(May, 1969), 164-168.
- M.J. Rosenberg, C.I. Hovland, W.J. McGuire, R.P. Abelson and W.J. Brehm, Attitude Organization and Change (New Haven, Conn.: Yale University Press, 1960)
- D.O. Sears, "Political Behavior," in G. Linzey and E. Aronson (eds.), Handbook of Social Psychology, V (Reading, Mass.: Addison-Wesley, 1969), 315-457.
- N. Schmitt, B.W. Coyle and B.B. Saari, "A Review and Critique of Analyses of Multitrait-Multimethod Matrices" Multivariate Behavioral Research, 12(December, 1977), 447-478.
- A. Wicker, "Attitudes Versus Actions: The Relationship of Verbal and Overt Behavioral Response to Attitude Objects," Journal of Social Issues, 25(1969), 41-47.
- G. Zaltman, C.R.A. Angelmar and R. Angelmar, Metatheory and Consumer Research (New York: Holt, Rinehart and Winston, 1973).

ATTITUDE MEASUREMENT AND BEHAVIOR CHANGE: A RECONSIDERATION OF  
ATTITUDE ORGANIZATION AND ITS RELATIONSHIP TO BEHAVIOR

Richard P. Bagozzi, University of California, Berkeley  
Robert E. Burnkrant, Ohio State University\*

ABSTRACT

The validity of a single component model of attitude is assessed and compared with a multi-component attitude conceptualization. Convergent and nomological validity are found for a two component conceptualization of attitude, while the single component model is not supported. The implications of these findings for consumer research are discussed.

Introduction

A considerable body of consumer behavior literature has dealt with the attitude-behavior relationship. Almost invariably attitude has been treated as a function of beliefs and their associated values (i.e., Hansen, 1969, Bass and Talarzyk, 1972; Kraft, Granbois and Summers, 1973; Lutz, 1975). These attitudes have been related to behavioral criteria with mixed but frequently disappointing results. It has been pointed out that some of the weaker results may have been due to the use of improper measurement procedures (i.e., Cohen, Fishbein and Ahtola, 1972; Cohen, 1972) or to a lack of correspondence between attitudinal predictors and behavioral criteria (i.e., Ajzen and Fishbein, 1977). The more recent use of attitude toward an act to predict specific intentions or behaviors reflects this concern for correspondence between attitude and behavior (i.e., Ryan and Bonfield, 1975; Wilson, Mathews and Harvey, 1975; Lutz, 1977).

The earlier research applying attitude to consumer behavior was concerned primarily with demonstrating the predictive validity of the attitude construct (cf., Wilkie and Pessemier, 1973). More recently it has been demonstrated that changes in beliefs lead to changes in attitude (Lutz, 1975) and that people combine belief and evaluation multiplicatively in their formation of an attitude (Bettman, Capon and Lutz, 1975). These contributions have undoubtedly advanced our understanding of attitudes and their application to consumer behavior. It is our contention, however, that the basic issue of attitude organization and its consistency with attitude measurement procedures has been largely ignored in this literature.

Two major distinguishable conceptualizations of attitude have flourished in the psychology literature. Those who hold a multi-dimensional view consider attitude to be a complex construct comprised of two or more components. In accordance with this view, Krech, Crutchfield and Ballachey (1962) define attitude as an enduring system of cognitions, feelings and response dispositions centered about a single object. Similarly, Rosenberg and Hovland (1960) regard attitude as a predisposition to some class of stimuli with cognitive, affective and behavioral responses. It is implied by those who hold a multi-component view of attitude that this attitude-behavior relationship would be stronger when the components are consistent than when they are inconsistent (Rosenberg and Hovland, 1960; Rosenberg, 1968). The multi-component view suggests that cognitive, affective and conative evaluations of objects are distinguishable aspects of attitude and that simultaneous consideration of all three components should be most predictive of overt behavior (cf., Greenwald, 1968). Failure to find a consistent direct relationship between attitude and behavior may be due to a failure to measure people's standing on all three components of attitude and to employ

these as simultaneous and/or independent predictors of behavior.

The major alternative to the three component view treats attitude as a single affective construct. For example, Thurstone defines attitude as "the affect for or against a psychological object" (1931, p. 261). Fishbein (1967) argues that all attitude scaling techniques have in common the characteristic that they place individuals on a dimension of affect. This affect for or against an object is typically inferred from an assessment of people's beliefs about the object and the evaluative aspect of those beliefs. Therefore, alternative approaches to the measurement of attitude provide alternative measures of the same thing (i.e., affect) and should yield the same results. Obtained differences among alternative instruments in measurement of an attitude, according to this approach, would be due to measurement error and not the assessment of alternative components.

A third and intermediate position is maintained in the present paper. According to this view, attitude is a complex construct comprised of cognitive and affective components. These components simultaneously account for behavioral intentions. These intentions, in turn, lead to overt behaviors.

Katz and Stotland (1959) and Rosenberg (1968) point out that all true attitudes must have both cognitive and affective content, although they need not include a conative component. Similarly, Rosenberg (1968) stresses that, with the exception of cognitive dissonance, most of the consistency theories give only token recognition to the definition of attitude as an internally consistent structure of affective, cognitive and behavioral components; but, in practice, the behavioral component is usually treated as a dependent variable. The two component attitude position taken here recognizes and is consistent with the fact that self-reported behaviors and stated intentions to respond have frequently been treated as dependent effects of affective and/or cognitive variables (e.g., Tittle and Hill, 1967; Warner and DeFleur, 1969; Rogers and Thistlethwaite, 1970). Intentions appear to be at a lower level of abstraction (i.e., closer to observable behavior) than the cognitions and affective feelings on which they are based. We, therefore, propose that attitude be viewed as a two component construct comprised of a cognitive and an affective component. These two attitudinal dimensions are believed to simultaneously account for behavioral predispositions, although they may have a differential impact on them. Behavioral predispositions, in turn, lead to overt behaviors. The purpose of this paper will be to test and contrast this two component view to the single component model.

Attitude research in marketing, in contrast to this variation in conceptualizations, has almost invariably assumed that attitude is an undimensional construct representing the affect for or against a psychological object, event, or situation. This assumption is implicit in the operationalization of Fishbein's model. It is consistent with the fact that most studies in marketing obtain only one measure of attitude which they relate to behavior (Hansen, 1969). When multiple

methods are employed to measure attitude they are treated as alternative measures of the same thing (i.e., attitude) rather than measures of alternative components of attitude (e.g., Lutz, 1975; Lutz, 1977).

Several attempts have been made to provide empirical support for a multi-component treatment of attitude, but they have yielded mixed results. Woodmansee and Cook (1967) failed to find components which could be interpreted as representing affective, cognitive and conative components of attitude when they factor analyzed scales designed to measure attitudes toward blacks. Ostrom (1969) and Kothandapani (1971) employed judgment procedures to develop Guttman, Likert, Thurstone and Guilford self-rating scales for each of the three components. While both researchers claim that their results support a three component model of attitude, Bagozzi's (1978) reanalysis of their results obtained support for a multi-component attitude only with Ostrom's data.

The present research involves a reanalysis of data initially presented by Fishbein and Ajzen (1974). In that study, each subject's attitude toward "being religious" was assessed on Thurstone, Guttman, Likert, semantic differential and Guilford self-rating scales. In addition, subjects provided responses to either 100 behavioral intention items or 100 self-reported behaviors. While Fishbein and Ajzen were consistent with a unidimensional approach to attitude in their treatment of all five attitude scales as alternative measures of attitude, it is our contention that the evaluative dimension of the semantic differential and the Guilford self-rating scale represent alternative measures of the affective component of attitude and that the Guttman, Likert and Thurstone scales represent alternative measures of the cognitive component of attitude.

Osgood, Suci and Tannenbaum (1957) define attitude as the projection of a concept on the evaluative dimension of semantic space. This unidimensional definition identifies attitude as the single dimension of semantic space which accounts for the concept's goodness or badness. Katz and Stotland (1959) identify the affective component with attributions of good or bad qualities. Similarly, McGuire (1969) argues that the evaluative dimension of semantic space is a measure of affect. Norman (1975) recently treated the evaluative dimension of semantic space as an operationalization of the affective component of attitude.

The Guilford self-rating scale asked people to provide their attitude toward being religious by checking an 11 point scale ranging from *extremely favorable* to *extremely unfavorable*. According to Ostrom (1969, p. 16) the affective component contains "statements representing favorable or unfavorable feelings." Kothandapani (1971, p. 323) made the same statement in his definition of the affective component. Similarly, Norman (1975) used a direct rating of a concept's favorability as an operationalization of the affective component of attitude toward that concept.

"The other three attitude measures were standard religiosity scales based on opinion items" (Fishbein and Ajzen, 1974, p. 62). Elsewhere, Fishbein and Ajzen (1975), p. 12) consider opinions to be cognitions or beliefs, and they distinguish them from affect and conation. The standard scales Fishbein and Ajzen (1974) employed included a Likert scale from Bardis (1961), a Guttman scale from Faulkner and De Jong (1969), and a Thurstone scale from Poppleton and Pilkington (1963). The statements employed in these scales are of the following form:

1. "Religious faith is merely another name for belief which is contrary to reason" (Poppleton and Pilkington, 1963, p. 34).

2. "Religious truth is higher than any other form of trust" (Faulker and De Jong, 1969, p. 567).

The cognitive component of attitude accounts for the perceived relationships between the object of attitude and other objects or concepts. Ostrom points out that the cognitive component of attitude includes, "beliefs about the object, characteristics of the object, and relationships of the object with other objects" (1969, p. 16). Kothandapani (1971) notes that this component includes beliefs and opinions about the object. Examination of the Guttman, Likert, and Thurstone scales employed by Fishbein and Ajzen shows them to be consistent with this conceptualization of the cognitive component of attitude.

It is possible to distinguish empirically between the single component and the two component view of attitude using Fishbein and Ajzen's data. A single component model of attitude would be supported if convergent validity is obtained when all five attitude scales are treated as alternative measures of attitude. A two component attitude model would be supported if convergence is obtained only when semantic differential and Guilford self-rating scales are treated as alternative measures of the affective component; and Guttman, Likert, and Thurstone scales are treated as alternative measures of the cognitive component of attitude. Furthermore, the two component position requires that both the cognitive and affective components simultaneously account for behavioral intentions and self-reported behaviors. The data do not permit us to distinguish between the prediction that behavior is accounted for by intentions alone and the prediction that behavior is accounted for by the simultaneous consideration of the conative, cognitive, and affective components of attitude. This is due to the fact that behavior and behavioral intention were not both obtained from the same individuals. However, a finding that the cognitive and affective components of attitude simultaneously account for behavioral intention and verbal reports of past behavior would be consistent with the two component view suggested here.

In summary, we expect that a reanalysis of the Fishbein and Ajzen data will obtain convergent validity when the cognitive and affective measures of attitude are treated as separate components, but convergence will not be obtained when all five instruments are treated as alternative measures of the same underlying construct. Furthermore, we expect that each component of attitude will separately account for scaled multiple act criteria, and both components of attitude will simultaneously account for scaled multiple act criteria.

#### Methodology

#### Subjects and Measures

Two samples of respondents were obtained by Fishbein and Ajzen (1974). In the first, 62 male and female undergraduates indicated which items from a set of 100 behaviors they had performed (the self-reported behaviors sample). The set of behaviors consisted of a list of 70 actions dealing with religious matters (e.g., pray before or after meals, donate money to a religious institution) and 30 additional actions in a refusal format selected from the original 70 (e.g., refuse to state religious preference during university registration). The second sample was composed of 63 male and female undergraduates who indicated which items from the set of 100 behaviors they would perform (the behavioral intentions sample).

In addition, all subjects completed five scales measuring attitudes toward religion. A Guilford self-rating scale measured attitudes towards being religious on an



11-point scale ranging from extremely favorable to extremely unfavorable. A semantic differential scale measured evaluations of "being religious" on five 11-point bipolar scales having the following end points: good-bad, harmful-beneficial, wise-foolish, pleasant-unpleasant, and sick-healthy. The Guilford and semantic differential scales are considered in the present study to tap largely the evaluative or affective dimension of attitudes. Support for this contention was provided in the preceding sections of this article.

To measure the cognitive component of attitude toward religion, three standard religiosity scales were employed. These include Likert (Bardis, 1961), Guttman (Falkner & De Jong, 1969), and Thurstone (Popleton & Pilkington, 1963) scales.

Method of Analysis

*Convergent Validity.* Before testing the hypotheses relating attitudes to behavior, it is necessary to establish the validity of the attitude measures.<sup>1</sup> Figure 1 presents two path diagrams that can be used in this regard. The single factor model of Figure 1a hypothesizes that the five attitude measures (i.e., the Guilford self-report, SR; semantic differential, SD; Guttman, G; Likert L; and Thurstone T; scales) each indicate a single underlying construct, "attitudes" (i.e.,  $A_{1+2}$ ). Fishbein and Ajzen (1974) did not differentiate the five attitude measures into affective and cognitive dimensions. Rather, the authors treated all five as independent measures of the same underlying "attitude towards religion" construct. Thus, they were implicitly assuming that the scales converged to measure a single attitude construct. This assumption is represented in Figure 1a. The validity of this assumption as well as the bicomponent model will be tested explicitly in the present study before investigating the attitude-behavior hypotheses (see below).

The two factor model of Figure 1b posits that attitudes are represented as two conceptually independent, yet empirically related constructs: (1) an affect or evaluative component ( $A_1$ ) and (2) a cognitive or belief dimension ( $A_2$ ).

To test the models of Figure 1, a confirmatory factor analysis methodology can be used (cf., Kenny, 1976). Briefly, the general confirmatory factor analysis model to test is

$$y = \Lambda x + z \tag{1}$$

where  $y$  is a vector of  $p$  measurements,  $x$  is a  $k < p$  vector of factors,  $z$  is a vector of  $p$  unique scores, and  $\Lambda$  is a  $p \times k$  matrix of factor loadings. For Figure 1a,  $p = 5$  and  $k = 1$ ; while for Figure 1b,  $p = 5$  and  $k = 2$ . With the assumptions that  $E(x) = E(z) = 0$ ,  $E(xx') = \Phi$ , and  $E(zz') = \Psi$ , where  $\Psi$  is a diagonal matrix, the variance-covariance matrix of  $y$  may be expressed as

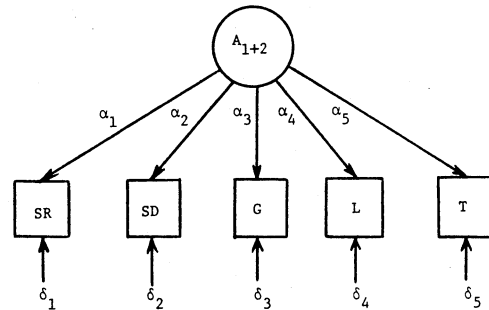
$$\Sigma = \Lambda\Phi\Lambda' + \Psi \tag{2}$$

Jöreskog (1969) derives a maximum likelihood procedure for estimating the parameters in  $\Lambda$ ,  $\Phi$ , and  $\Psi$ . Further, the methodology yields an overall  $\chi^2$  goodness-of-fit

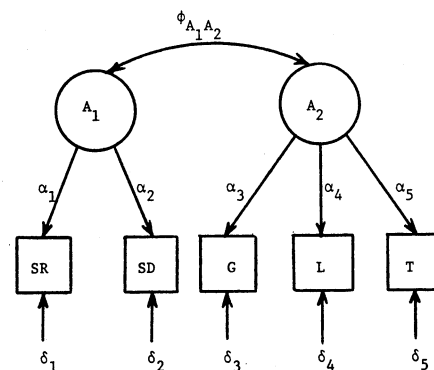
<sup>1</sup>Although a full test of construct validity would require the use of the multitrait-multimethod research design, this was not possible in the present study because the same multiple measures were not used on each of the two hypothesized attitudinal components. However, because two measures of affect and three measures of beliefs were obtained, it is possible to ascertain the degree of convergent validity for the attitude measures. Further, as additional evidence for construct validity, the data allow also for the determination of nomological validity as developed below.

test. The computer program, LISREL, may be used to test the models of Figure 1 and equations (1) and (2) (Jöreskog & van Thillo, 1972).

FIGURE 1  
PATH DIAGRAMS FOR DETERMINING CONVERGENT VALIDITY OF ATTITUDE MEASURES



A. Single Factor Model



B. Two Factor Model

*Nomological Validity.* As further evidence of construct validity, the degree of nomological validity of the attitude measures was determined. Following Campbell (1960), nomological validity is defined herein as the degree to which predictions from a concept in a theoretical system of concepts are confirmed. Figure 2 presents a path model that can be used to test for the nomological validity of attitudes where, for purposes of illustration, attitudes ( $A$ ) are shown measured by three scales ( $y_1, y_2,$  and  $y_3$ ) and behavior ( $B$ ) is indicated by two scales ( $y_4$  and  $y_5$ ). The hypothesis is that the nomological validity of the attitude construct ( $A$ ) will be confirmed if predictions from the model as to the resulting behavior ( $B$ ) are validated. The proper attitude construct ( $A$ ) to use is the one achieving convergent validity as determined in the previous analysis. The appropriate behavioral measure ( $B$ ) is the one constructed from scaled behaviors because these behaviors approach the same level of specificity as the attitudinal items. For a description of the scaling procedures used, the reader is referred to Fishbein and Ajzen (1974).

The hypothesis of nomological validity can be tested using Jöreskog's analysis of covariance structures model (Jöreskog, 1970). That is, the model to test is

$$B = \gamma A + \zeta \tag{3}$$

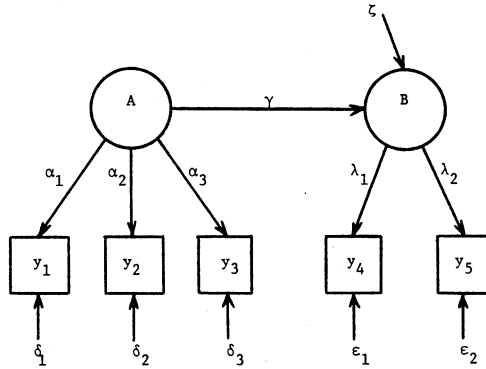
where it is assumed that all variables are taken to have zero expectations and that the disturbances (i.e.,  $\zeta, \delta_1, \delta_2, \epsilon_1,$  and  $\epsilon_2$ ) are mutually independent and are independent of their corresponding explanatory variables. The computer program, LISREL, can be used to

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix} = \begin{pmatrix} \alpha_1 & 0 \\ \alpha_2 & 0 \\ \alpha_3 & 0 \\ 0 & \lambda \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \epsilon_1 \\ \epsilon_2 \end{pmatrix} \quad (4)$$

estimate parameters and test the model of Figure 2 and equations (3) and (4). Although the model is illustrated for the case of three attitudinal and two behavioral measures, it can be generalized to the case of  $\lambda$  attitudinal and  $m$  behavioral measures.

FIGURE 2

PATH DIAGRAM FOR DETERMINING THE RELATIONSHIP BETWEEN ATTITUDES AND SCALED MEASURES OF BEHAVIOR



Results

Looking only at the intercorrelations of variables in both samples, Fishbein and Ajzen (1974) concluded that the five verbal attitude scales showed "a high degree of convergent validity." This and the other hypotheses discussed above will be investigated using the structural equation methodology.

Convergent Validity

The findings for the tests of convergent validity using the more rigorous structural equation methodology, are shown in Table 1. Notice first that the hypothesis of convergence for the single factor model must be rejected for both the self-reported behaviors sample (i.e.,  $\chi^2 = 23.91$ , d.f. = 5,  $p \approx .00$ ) and the behavioral intentions sample ( $\chi^2 = 13.98$ , d.f. = 5,  $p \approx .02$ ). Thus, contrary to the original claim made by Fishbein and Ajzen (1974) and contrary to the assumptions made in most consumer research, one can not accept the hypothesis that the five scales each measure a single underlying attitude construct.

Consequently, the two factor model of attitudes was tested. As shown in the final two columns of Table 1 the hypothesis of convergence receives strong support for the self-reported behaviors sample (i.e.,  $\chi^2 = 2.30$ , d.f. = 4,  $p \approx .68$ ) and adequate support for the behavioral intentions sample (i.e.,  $\chi^2 = 8.30$ , d.f. = 4,  $p \approx .08$ ).<sup>2</sup> Thus, convergent validity is established for the two factor affective/cognitive model of attitudes but not for the single factor model. Any further analysis must take this finding into consideration.

<sup>2</sup>Although the  $\chi^2$  value for the behavioral intentions sample indicates a borderline fit, (i.e.,  $p \approx .08$ ), when Bartlett's (1951) small sample correction factor is applied, the model reaches acceptable levels of significance (i.e.,  $\chi^2 = 7.65$ , d.f. = 4, and  $p \approx .10$ ).

TABLE 1

GOODNESS-OF-FIT TESTS AND PARAMETER ESTIMATES FOR CONSTRUCT VALIDITY MODELS

Parameter	Single Factor Model		Two Factor Model	
	Self-reported Behaviors	Behavioral Intentions	Self-reported Behaviors	Behavioral Intentions
$\alpha_1$	.736	.789	.822	.831
$\alpha_2$	.840	.866	.973	.920
$\alpha_3$	.828	.919	.844	.925
$\alpha_4$	.924	.953	.934	.954
$\alpha_5$	.847	.836	.852	.840
$\theta_{\delta_1}$	.677	.614	.569	.556
$\theta_{\delta_2}$	.543	.501	.231	.391
$\theta_{\delta_3}$	.561	.394	.536	.380
$\theta_{\delta_4}$	.385	.302	.356	.300
$\theta_{\delta_5}$	.532	.548	.524	.543
$\phi_{A_1 A_2}$	-	-	.824	.927
Tests of Models	$\chi^2 = 23.91$ d.f. = 5 $p \approx .00$	$\chi^2 = 13.98$ d.f. = 5 $p \approx .02$	$\chi^2 = 2.30$ d.f. = 4 $p \approx .68$	$\chi^2 = 8.30$ d.f. = 4 $p \approx .08$

Nomological Validity

Given that two dimensions of attitude have been identified, nomological validity was determined for each. Looking first at the relation between the affective component of attitude and the scaled behavior measures, it can be seen in Table 2 that nomological validity is established for both the self-reported behaviors and behavioral intentions samples. This occurs in either case whether behavior is measured with the Guttman and Likert techniques or the Likert and Thurstone techniques.<sup>3</sup> Also, as illustrated by the standardized estimates for  $\gamma$ , the evaluative attitudes toward the church relate at a high level of magnitude to the scaled behavior measures, as predicted by theory.

Looking next at the relation between the cognitive component of attitude and the scaled behavior measures, it can be seen in Table 3 that nomological validity is again established for both samples and for both pairs of scaled behavior measures. Further, the cognitive components relate at a high level of magnitude to the scaled behavior measures, as predicted by theory.

In sum, based on the results for convergent and nomological validity, some evidence exists for establishing the construct validity of the affective/cognitive model of attitudes.

To investigate the differential effects of affect and cognitions on behavior, the path model of Figure 3 may be used. The model to test is based on Jöreskog's analysis of covariance structure:

<sup>3</sup>Although the  $\chi^2$  value for the behavioral intentions sample indicates a borderline fit, when Bartlett's (1951) small sample correction factor is applied, the model reaches acceptable levels of significance.

TABLE 2

GOODNESS-OF-FIT TESTS AND PARAMETER ESTIMATES FOR NOMOLOGICAL VALIDITY MODELS RELATING AFFECT TO BEHAVIOR

Parameter	Self-reported Behaviors		Behavioral Intentions	
	G&L	L&T	G&L	L&T
$\alpha_1$	.816	.863	.843	.846
$\alpha_2$	.980	.927	.908	.904
$\lambda_1$	.813	.848	.851	.941
$\lambda_2$	.954	.934	1.000	.843
$\gamma$	.737	.849	.705	.750
$\theta_{\delta_1}$	.578	.505	.539	.533
$\theta_{\delta_2}$	.197	.375	.419	.428
$\theta_{\epsilon_1}$	.582	.530	.525	.339
$\theta_{\epsilon_2}$	.300	.357	.000	.538
$\psi$	.458	.279	.503	.437
Tests of Models	$\chi^2 = 1.01$	$\chi^2 = 2.77$	$\chi^2 = 1.38$	$\chi^2 = .14$
	d.f. = 1	d.f. = 1	d.f. = 1	d.f. = 1
	p $\approx$ .32	p $\approx$ .10	p $\approx$ .24	p $\approx$ .71

TABLE 3

GOODNESS-OF-FIT TESTS AND PARAMETER ESTIMATES FOR NOMOLOGICAL VALIDITY MODELS RELATING COGNITIONS TO BEHAVIOR

Parameter	Self-reported Behaviors		Behavioral Intentions	
	G&L	L&T	G&L	L&T
$\alpha_1$	.861	.869	.935	.938
$\alpha_2$	.918	.911	.938	.935
$\alpha_3$	.857	.858	.857	.859
$\lambda_1$	.852	.938	.851	.904
$\lambda_2$	.911	.844	1.000	.877
$\gamma$	.798	.777	.742	.819
$\theta_{\delta_1}$	.508	.494	.355	.348
$\theta_{\delta_2}$	.396	.412	.346	.355
$\theta_{\delta_3}$	.515	.514	.515	.513
$\theta_{\epsilon_1}$	.524	.347	.525	.427
$\theta_{\epsilon_2}$	.412	.536	.000	.481
$\psi$	.363	.396	.449	.329
Tests of Models	$\chi^2 = 5.38$	$\chi^2 = .85$	$\chi^2 = 6.35$	$\chi^2 = 4.79$
	d.f. = 4	d.f. = 4	d.f. = 4	d.f. = 4
	p $\approx$ .25	p $\approx$ .93	p $\approx$ .17	p $\approx$ .31

$$B = (\gamma_1 \gamma_2) \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} + \zeta \quad (5)$$

$$\begin{pmatrix} L \\ T \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} B + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix} \quad (6)$$

$$\begin{pmatrix} SR \\ SD \\ G \\ L \\ T \end{pmatrix} = \begin{pmatrix} \alpha_1 & 0 \\ \alpha_2 & 0 \\ 0 & \alpha_3 \\ 0 & \alpha_4 \\ 0 & \alpha_5 \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \end{pmatrix} \quad (7)$$

where it is assumed that all variables are taken to have zero expectations and that the disturbances are mutually independent and are independent of their corresponding explanatory variables. Again, LISREL may be used to estimate the parameters in the model and test hypotheses.

Applying LISREL to the data yields the results shown in Table 4 where only the findings for the self-reported behaviors sample are displayed. In general, the model fits the data well ( $\chi^2 = 13.11$ , d.f. = 10,  $p \approx .22$ ). Further, the affective or evaluative component can be seen to be approximately three times as forceful in its impact on behavior as the cognitive component<sup>4</sup> ( $\gamma_1 = .651$  vs  $\gamma_2 = .226$ ). It should be noted, however, that the values for  $\gamma_1$  and  $\gamma_2$  will be sensitive to the pattern of correlations among the behavior and attitude measures, and problems of multicollinearity should be considered.

#### Discussion

The results obtained from a reanalysis of the Fishbein and Ajzen data support a two component model of attitude. When Guilford self-rating, semantic differential, Guttman, Likert and Thurstone scales were treated as alternative measures of a single underlying construct (i.e., attitude), the analysis failed to provide convergent validity. When Guilford self-rating and semantic differential scales were treated as alternative measures of an affective component of attitude and Guttman, Likert and Thurstone scales were treated as alternative measures of a cognitive component of attitude, convergent validity was supported. The validity of the cognitive-affective model was supported for both the self-report and the behavioral intentions samples.

Further support for the validity of the two component model is provided by a consideration of nomological validity. It was found that each component of the two component model separately predicted scaled multiple act criteria. Furthermore, both components contributed to the simultaneous prediction of behavior. While both components simultaneously accounted for behavior, the affective component was approximately three times as powerful as the cognitive component. In summary, then, the results provide support for the nomological validity of the two component attitude model. Behavior was predicted from a simultaneous consideration of the cognitive and affective components of attitude.

These results and those obtained from Bagozzi's (1978) reanalysis of Ostrom's (1969) data provide strong

<sup>4</sup>The nomological validity models are run with the Guttman/Likert and Likert/Thurstone operationalizations rather than with a single Guttman/Likert/Thurstone operationalization of behavior because inspection of the correlation matrix in Table 1 reveals that the Guttman and Thurstone measures correlate at a lower level than the other two pairs in both samples.

FIGURE 3  
THE IMPACT OF AFFECT AND COGNITIONS ON BEHAVIOR

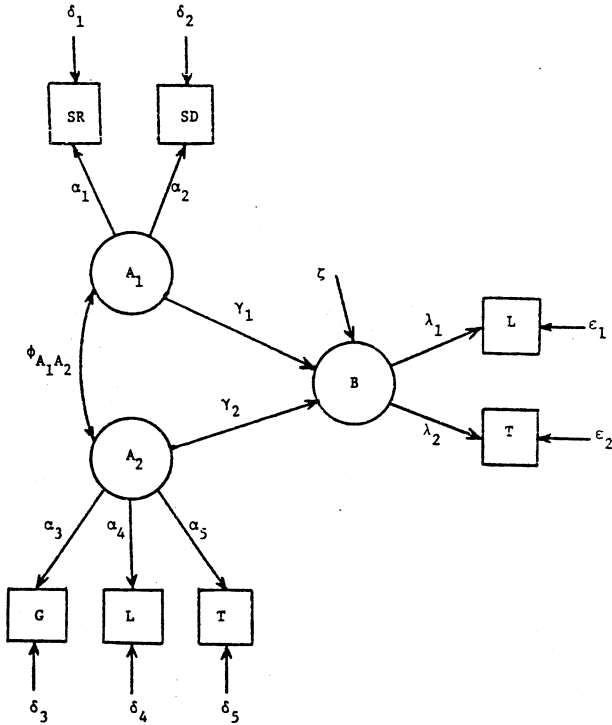


TABLE 4

GOODNESS OF FIT TESTS AND PARAMETER ESTIMATES FOR A MODEL TESTING THE RELATIVE IMPACT OF AFFECT AND COGNITIONS ON BEHAVIOR

Parameter	Self-reported Behaviors	
	L	T
$\alpha_1$	.838	
$\alpha_2$	.955	
$\alpha_3$	.849	
$\alpha_4$	.930	
$\alpha_5$	.853	
$\gamma_1$	.651	
$\gamma_2$	.226	
$\lambda_1$	.879	
$\lambda_2$	.901	
$\theta_{\delta_1}$	.546	
$\theta_{\delta_2}$	.296	
$\theta_{\delta_3}$	.528	
$\theta_{\delta_4}$	.367	
$\theta_{\delta_5}$	.522	
$\theta_{\epsilon_1}$	.477	
$\theta_{\epsilon_2}$	.434	
$\phi_{A_1 A_2}$	.834	
$\psi$	.279	
Tests of Model	$\chi^2$	= 13.11
	d.f.	= 10
	p	≈ .22

support for a multi-component treatment of attitude. Attitude is regarded as a complex construct comprised of affective and cognitive components. These components were found to account for behavior and behavioral intentions.

It was pointed out earlier that most attitude research in consumer behavior has been based on measurement of only one component of attitude. Both in research dealing with attitude change and in research considering the attitude-behavior relationship, conclusions are generally made after measuring either the cognitive or the affective component of attitude. It has been maintained that failure to obtain consistent results in research dealing with the attitude-behavior relationship may be due, in part, to the failure to measure more than one component of attitude in any given study. The results obtained from the present research are consistent with the argument that a complete accounting of attitude requires measurement of both the cognitive and affective components.

Greenwald (1968) suggests that the cognitive and affective components of attitude have distinct antecedents. The affective component may be formed most directly through classical conditioning whereas the cognitive component may be more directly affected by communication and cognitive learning. This indicates that attitude change strategies may have differential effects on the components of attitude. Even though a message is quite effective in changing the cognitive component of attitude it may have a relatively weak effect on behavior if the affective component is a more powerful predictor of the given behavior. This issue has received little attention although it may have important implications for the effects of attitude change strategies on behavior change.

In this study, even though both components of attitude were found to simultaneously account for behavior, the affective component was found to be roughly three times as forceful in its impact on behavior as the cognitive component. It may be that the relative impact of the various components of attitude is dependent on the attitude-behavior target under consideration. We may speculate that the affective component may play a stronger role in the purchase of low involvement products such as soft drinks, for instance, whereas the cognitive component may play a larger role in the purchase of products requiring a more complex decision process prior to purchase. The findings of the present study may tend to argue against this since we may expect religion to be highly involving, although the performance of religious behaviors may not involve a very conscious or extended decision process. The point is, however, that both components should be investigated as simultaneous predictors of purchase behavior. The relative impact of the cognitive and affective components may vary in terms of the nature of the product under consideration.

In the Fishbein and Ajzen study the cognitive and affective measures of attitude were highly correlated. This may be due to pressure toward cognitive consistency. Even though the cognitive and affective components of attitude may be formed by distinct processes, the pressure to achieve affective-cognitive consistency probably leads to high correlations among the components of attitude. As McGuire (1960) has pointed out, consistency is likely to increase over time. The relatively weak relationship between cognitive and affective components of attitude shown in the Lutz (1977) study may be due to the fact that, in contrast to Fishbein and Ajzen, Lutz manipulated beliefs and then measured the effects of this manipulation on cognitive and affective measures of attitude. It may be that a stronger relationship would have been obtained if a delayed post-

test had been employed. The delay would have provided more time for consistency restoration to take effect.

In summary, the results of this study show that attitude is comprised of two distinct components. These components simultaneously account for behavioral criteria. Proper assessment of the attitude behavior relationship, therefore, requires measurement and use of both of these attitude components as simultaneous predictors of behavior.

#### References

\* Both authors contributed equally to the research.

I. Ajzen and M. Fishbein, "Attitude-Behavior Relations: A Theoretical Analysis and Review of Empirical Research," Psychological Bulletin, 84 (1977), 888-918.

R. P. Bagozzi, "The Construct Validity of the Affective, Behavioral, and Cognitive Components of Attitude by Analysis of Covariance Structures," Multivariate Behavioral Research, 13 (1978), 9-31.

P. D. Bardis, "A Religion Scale," Social Science, 36 (1961), 120-123.

M. S. Bartlett, "The Effect of Standardization on an Approximation in Factor Analysis," Biometrika, 38 (1951), 337-344.

F. M. Bass and W. W. Talarsyk, "An Attitude Model for the Study of Brand Preference," Journal of Marketing Research, 9 (1972), 93-96.

J. R. Bettman, N. Capon, and R. J. Lutz, "Multiattribute Measurement Models and Multiattribute Attitude Theory: A Test of Construct Validity," Journal of Consumer Research, 1 (1975), 11-15.

D. T. Campbell, "Recommendations for APA Test Standards Regarding Construct, Trait, or Discriminant Validity," American Psychologist, 15 (1960), 546-553.

J. B. Cohen, "Toward an Integrated Use of Expectancy-Value Attitude Models." Paper presented at the ACRAMA workshop on buyer/consumer information processing, Chicago (1972).

J. B. Cohen, M. Fishbein, and O. T. Ahtola, "The Nature and Uses of Expectancy-Value Models in Consumer Attitude Research," Journal of Marketing Research, 9 (1972), 456-460.

J. E. Faulker and G. F. De Jong, "Religiosity in 5-D: An Empirical Analysis." In J. P. Robinson and P. R. Shaver (eds.), Measures of Social Psychological Attitudes. Ann Arbor: University of Michigan, Institute of Social Research (1969).

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).

M. Fishbein and I. Ajzen, "Attitude Towards Objects as Predictors of Single and Multiple Behavioral Criteria," Psychological Review, 81, (1974), 59-74.

M. Fishbein and I. Ajzen, Belief, Attitude Intention and Behavior: An Introduction to Theory and Research. Reading, Mass. Addison, Wesley (1975).

A. G. Greenwald, "On Defining Attitude and Attitude Theory." In A. G. Greenwald, T. C. Brock and T. M. Ostrom (eds.), Psychological Foundations of Attitudes. New York: Academic Press (1968).

F. Hansen, "Consumer Choice Behavior: An Experimental Approach," Journal of Marketing Research, 6 (1969) 436-443.

K. G. Jöreskog, "A General Approach to Confirmatory Maximum Likelihood Factor Analysis," Psychometrika, 34 (1969), 183-202

K. G. Jöreskog, "A General Method for Analysis of Covariance Structures," Biometrika, 57 (1970), 239-251.

K. G. Jöreskog and M. van Thillo, LISREL: A general computer program for estimating a linear structural equation system involving multiple indicators of unmeasured variables. Educational Testing Service Research Bulletin, (1972), 72-156.

D. Katz and E. Stotland, "A Preliminary Statement to a Theory of Attitude Structure and Change." In S. Koch (ed.), Psychology: A Study of a Science, Vol 3, New York: McGraw-Hill (1959), 423-475.

D. A. Kenny, "An Empirical Application of Confirmatory Factor Analysis to the Multitrait-Multimethod Matrix," Journal of Experimental Social Psychology, 12 (1976), 247-252.

V. Kothandapani, "Validation of Feeling, Belief, and Intention to Act as Three Components of Attitude and Their Contribution to Prediction of Contraceptive Behavior," Journal of Personality and Social Psychology, 19 (1971), 321-333.

F. B. Kraft, D. H. Granbois and J. O. Summers, "Brand Evaluation and Brand Choice: A Longitudinal Study," Journal of Marketing Research, 10 (1973), 235-241.

D. Krech, R. S. Crutchfield and E. L. Ballachey, Individuals in Society. New York: McGraw-Hill (1962).

R. J. Lutz, "Changing Brand Attitudes Through Modification of Cognitive Structure," Journal of Consumer Research, 1, (1975), 49-59.

R. J. Lutz, "An Experimental Investigation of Causal Relations Among Cognitions, Affect, and Behavioral Intention," Journal of Consumer Research, 3 (1977), 197-208.

W. J. McGuire, "Cognitive Consistency and Attitude Change," Journal of Abnormal and Social Psychology, 60 (1960), 345-353.

W. J. McGuire, "The Nature of Attitudes and Attitude Change." In G. Lindzey and E. Aronson (eds.), The Handbook of Social Psychology, Vol. 3, Cambridge, Mass: Addison-Wesley (1969).

R. Norman, "Affective-Cognitive Consistency, Attitudes, Conformity, and Behavior," Journal of Personality and Social Psychology, 32 (1975), 83-91.

C. E. Osgood, G. J. Suci and P. H. Tannenbaum, The Measurement of Meaning. Urbana: University of Illinois Press (1957).

T. M. Ostrom, "The Relationship Between the Affective, Behavioral, and Cognitive Components of Attitude," Journal of Experimental Social Psychology, 5 (1969), 12-30.

P. K. Poppleton and G. W. Pilkington, "The Measurement of Religious Attitudes in a University Population," British Journal of Social and Clinical Psychology, 2 (1963), 20-36.

R. W. Rogers and D. L. Thistlethwaite, "Effects of Fear Arousal and Reassurance on Attitude Change," Journal of

Personality and Social Psychology, 15 (1970), 227-233.

M. J. Rosenberg and C. I. Hovland, "Cognitive, Affective and Behavioral Components of Attitudes." In M. J. Rosenberg, C. I. Hovland (eds.), Attitude Organization and Change: An Analysis of Consistency Among Attitude Components. New Haven: Yale University Press (1960).

M. J. Rosenberg, "Hedonism, Inauthenticity, and Other Goals Toward Expansion of a Consistency Theory." In R. P. Abelson, et al. (eds.), Theories of Cognitive Consistency: A Source-Book. Chicago: Rand McNally (1968).

M. J. Ryan and E. H. Bonfield, "The Fishbein Extended Model and Consumer Behavior," Journal of Consumer Research, 2 (1975), 118-136.

L. L. Thurstone, "The Measurement of Social Attitudes," Journal of Abnormal and Social Psychology, 26, (1931), 249-269.

C. R. Tittle and R. J. Hill, "Attitude Measurement and Prediction of Behavior: An Evaluation of Conditions and Measurement Techniques," Sociometry, 30 (1967), 199-213.

L. G. Warner and M. L. De Fleur, "Attitude as an Interactional Concept: Social Constraint and Social Distance as Intervening Variables Between Attitudes and Action," American Psychological Review, 34, (1969), 153-169.

W. L. Wilkie and E. A. Pessemier, "Issues in Marketing's Use of Multiattribute Attitude Models," Journal of Marketing Research, 10, (1973), 428-441.

D. T. Wilson, H. L. Mathews, and J. W. Harvey, "An Empirical Test of the Fishbein Behavioral Intention Model." Journal of Consumer Research, 1 (1975), 39-48.

J. J. Woodmansee and S. W. Cook, "Dimensions of Verbal Racial Attitudes: Their Identification and Measurement." Journal of Personality and Social Psychology, 7 (1967), 240-250.

Joel B. Cohen, University of Florida\*

## Introduction

Two of the three papers presented in this session deal specifically with construct validity, and both of these choose the construct "attitude" as the basis for their discussion and empirical examination. The Cattin paper also addresses an important topic: cross validation of regression models to provide a more reliable estimate of fit to appropriate criterion variables. Because of the importance and particularly the underemphasis of the construct validity issue and the central role of attitudes in the study of consumer behavior, I have chosen to confine my comments to the former two papers and the rather similar positions they take.

## What is Construct Validity?

The 1974 revision of Standards for Educational and Psychological Tests defines a construct as, "an idea developed or 'constructed' as a work of informed, scientific imagination; that is...a theoretical idea developed to explain and to organize some aspects of existing knowledge." Central to the concept of construct validity is the dictum that a construct, "is a dimension understood or inferred from its network of interrelationships" (1974, p. 29).

In order to better understand the significance and implications of the above for empirical efforts at construct validation it is necessary to go back a few years. Between 1950 and 1954 a blue ribbon committee was convened by the American Psychological Association to establish standards for test validity. The preceding decade had seen much discussion and dissatisfaction with conventional notions of validity, and out of this came the committee's recommendation to distinguish four types of validity: predictive validity, concurrent validity, content validity, and, for the first time, construct validity. The first two of these are today often grouped together as criterion-related validity and apply when one wishes to infer an individual's standing or performance on some criterion variable from a test or trait score. What is of upmost concern here is that such a test "work" for the particular criterion variable of interest. Content validity is established by showing that the behaviors demonstrated in completing the test constitute a representative sample of behaviors to be found in the criterion domain of interest. Test items, for instance, must encompass all relevant facets (e.g., skills, knowledge) of a performance domain in order for the test to be a valid representation of that domain.

Construct validity rests on a philosophy of science position that, "to 'make clear what something is' means to set forth the laws in which it occurs...a nomological network" (Cronbach and Meehl, 1955, p. 290). The network generates testable propositions which relate test or trait scores (as representations of a construct) to other constructs some of which are observable through measurement or experimental observation. In a sense, the meaning of test or measure "validity" comes down to an issue of the specific types of inferences one wishes to make. A measure or test is merely a principle for making inferences (see Cronbach and Meehl, 1955, p. 297). Construct validity involves inferences to a network of related constructs and observables, the purpose being to better understand the nature of the construct being measured and to further develop the theoretical structure

in which it is imbedded.

## Campbell and Fiske's Approach

The construct validity position taken by Campbell and Fiske (1959), in setting forth the logic of the multi-trait-multimethod matrix, represents a concern with "the adequacy of tests as measures of a construct rather than with the adequacy of a construct as determined by the confirmation of theoretically predicted associations with measures of other constructs" (1959, p. 100). It is Campbell and Fiske's premise that, "before one can test the relationships between a specific trait and other traits, one must have some confidence in one's measures of that trait" (1959, p. 100).

It is this author's contention that the multitrait-multimethod approach is designed specifically to supply such confidence rather than as a paradigm for examining construct validity. In a subsequent, clarifying paper (Campbell, 1960) a distinction is drawn between "two types of construct validity": "trait validity" and "nomological validity." "Trait validity", to use Campbell's example, is illustrated by validating the Taylor Manifest Anxiety Scale against psychiatrists' ratings. "Nomological validity", on the other hand, is illustrated by interpreting such test scores within the theoretical structure of Hull-Spence learning theory (i.e., as a measure of D) and generating predictions of performance in learning situations.

Whenever possible, it makes a great deal of sense to compare a test with other independent means of measuring the same trait. It is also clear that such "trait validity" goes beyond criterion-based validity (i.e., no one test is the criterion), still the existence of higher correlations among alternative measures of the same trait relative to other traits may not go far in defining the nature of the construct or examining its role in a theoretical structure. It appears to be stretching things a bit, therefore, to think of "trait validity" as construct validity.

Much depends on the theoretical rationale for and nature of the traits and methods selected: subsequent inferences regarding validity must rest upon exactly what is being examined in the matrix. An example may clarify this. Campbell and Fiske term the monotrait-heteromethod values in their multitrait-multimethod matrix the "validity diagonal" and argue that entries in the validity diagonal provide evidence regarding convergent validity. Take the simplest case, in which each method of measurement for a particular trait is represented by a single item. According to Campbell and Fiske, convergent validity is supported by high correlations among these methods/items. The reader will recognize, however, that this is little more than high internal consistency reliability, which is itself, of course, an important aspect of test construction. In traditional multitrait-multimethod analyses, single items are replaced by entire scales. Still, the concept is the same--an assessment of the degree to which these measures "hang together." In discussing this issue, Campbell and Fiske argue that "reliability and validity can be seen as regions on a continuum," with test-retest and internal consistency reliability progressively closer to validity (1959, p. 83). The key issue to Campbell and Fiske lies in the

independence of measurement approaches: reliability representing convergence among maximally similar methods and validity representing convergence among maximally different methods.

Unless methods of measurement actually represent aspects of the nomological network, however, convergence among methods may do little to illuminate the construct. While methods qua nomological observables conceivably could fit this requirement (e.g., observation of performance theoretically predicted to covary with a trait or state measure), alternative test formats, objective vs. subjective means of assessment and the like cannot be said to be of this character.

Discriminant validity (as operationalized by Campbell and Fiske) would appear to be more directly relevant to construct validity provided that an absence of association among the traits selected for study is theoretically meaningful (i.e., enables the investigator to rule out certain interpretations regarding the meaning of the construct). In Campbell and Fiske's words, "One cannot define without implying distinctions, and the verification of these distinctions is an important part of the validation process" (1959, p. 84). For example, a lack of association between a test of aptitude and a test of specific and relevant knowledge enables the investigator to draw more precise inferences about the construct measured by the test. Or, an absence of systematic variance due to method-trait interaction may help the investigator draw the inference that the construct does not incorporate certain response set factors.

Selection of comparison traits and methods to fill out the matrix is, therefore, the key to any construct validity inferences which could emerge from the analysis. The role of each trait and method of measurement in the nomological network must be identified beforehand. Any interpretation of a multitrait-multimethod matrix with respect to construct validity (as opposed to confidence in the precision of one's measures) should, in this author's opinion, be carefully and critically examined. Convergent validity, in particular, may shed little light on the nature of the construct being studied.

Nothing said above, however, is intended to convey the impression that the multitrait-multimethod approach is anything other than an extremely important contribution to the logic and procedures of test and scale evaluation. Campbell and Fiske have made an extremely strong argument in support of multiple operationalism under the rubric of convergent validation: "Any single observation, as representative of concepts, is equivocal... the addition of a second viewpoint...greatly reduces this equivocality, greatly limits the constructs that could jointly account for both sets of data" (1959, p. 101). In addition, the procedures advocated by Campbell and Fiske provide a means of evaluating unwanted method variance, which may be an extremely important step in consumer research due to the customary use of pencil and paper instruments in standard response formats.

Establishing scale validity should not be thought of as a one-step task but an on-going process, and this is particularly true of construct validity. It is encouraging to see the greater concern the field is evidencing with respect to the definition and measurement of key constructs. Such concern appears on the threshold of being reflected in higher standards applied to measurement.

Both sets of authors follow the Campbell and Fiske approach and look upon the multitrait-multimethod approach as a test of construct validity. While nomological validity is mentioned in both papers, only the Bagozzi and Burnkrant paper actually attempts to specify part of the theoretical structure in which the construct "attitude" is embedded and to examine part of that structure.

What can we learn about construct validity from these papers? The John and Reve paper offers an assessment of alternative ways of analyzing data in a multitrait-multimethod matrix. This should be a helpful addition to methodological discussions seeking to refine analytical procedures in this area. The paper is limited by its substantive reliance on Ostrom's often analyzed data base, in the sense that any inferences to be made are constrained by the trait and method input to this matrix. There is no effort here to look at alternative conceptions of the attitude construct within a theoretical structure. The data, simply put, enable investigators to ask the question, "In the context of the methods used, to what degree are these components interrelated?" Making the remarkable assumption that each scale is a valid and reliable measure of the dimension it is intended to represent (i.e., either the cognitive, affective, or conative dimension), we are left without a means of evaluating the adequacy of any one or combination of these as a measure of attitude. Even if we were to conclude (contrary to the apparently more powerful analytical methods) that these were separately identifiable dimensions, on what basis do we conclude that each is a necessary component of attitude? Knowing that the dimensions are in some ways different from one another and can be identified as such using different methods of measurement is useful information, but what role, if any, each serves in identifying this construct and not something else is yet to be determined. One does not know that a valid measure of a construct exists until the scores obtained are shown to be consistent with predictions made about the covariation in observables tied by theory to the construct.

To clarify this, let's take a simple example. Say a theory is proposed linking two observables to a hypothetical construct. We'll term the observables S and R and the hypothetical construct O. Investigator 1 states that O is a unidimensional construct, while investigator 2 states that, in reality, O is made up of O<sub>1</sub>, O<sub>2</sub> and O<sub>3</sub>. To prove this the second investigator develops multiple means of measuring each component, administers the series of scales to a group of subjects and shows that each component is completely separate from the other two for each and every method used. Which investigator is right? Well, we've learned that the second investigator has three pretty good and distinct scales, but we really don't know whether he's done any better job representing the construct. That evidence awaits a study in which a prediction involving the construct and one or both observables is tested using such scores. The more complex the theoretical network, of course, the greater the amount and diversity of evidence that is required.

The Bagozzi and Burnkrant paper suffers from similar deficiencies. They propose a 2 component model of atti-

<sup>1</sup> Before going any further, I would like very much to applaud the impressive work by these four authors in carrying out an in-depth assessment of attitude measurement procedures. We need much more validity-oriented work in consumer behavior, and I believe it is high time the field raised its standards for both carrying out and reporting analyses of the validity and reliability of the measures relied on in research studies.



tudes and test this using data developed by Fishbein and Ajzen involving 5 different scaling methods, 2 of which appear to measure the affective and 3 of which the cognitive aspect of attitudes, together with a scaled behavioral intention measure (for part of the sample) and a scaled self-reported behavior measure (for the remaining subjects). Despite fairly reasonable intercorrelation among the five scales, the more rigorous structural equation technique used by the authors allows them to reject a single factor model in favor of their hypothesized 2 factor model, though the presence of an additional parameter in the two factor model needs somehow to be accounted for in evaluating the overall fit. To this point the authors claim only superior convergent validity for their model, stemming, it seems, from the (slightly) better fit with the measures partitioned into two groups.

Bagozzi and Burnkrant then propose to investigate the "nomological validity" of the two factor attitude measures, which (it has been argued earlier) is the essence of construct validity. Behavioral intention and self-reported behavior scores (approximately half the subjects for each) were used as the observables in their "nomological" network. Without question the concept behind this approach to construct validation is not only sound but represents a high water mark in consumer research. Procedurally, however, one must question the appropriateness of conceptualizing (to use Fishbein and Ajzen's words) "behavioral attitude scores" as a separate part of the nomological network. The three-component view of attitudes (as represented in the Ostrom data analyzed by John and Reve) suggests pushing behavioral intentions into the attitude construct itself, and the similarity among measurement methods (Guttman, Likert, Thurstone) used by Fishbein and Ajzen in developing scales for both attitudinal and behavioral elements may, in fact, tend to produce biases in the direction of consistency. I don't believe this is what Campbell and Fiske had in mind by "maximally different methods." In addition, even if maximally different, it is customary to treat both behavioral intentions and behavior as criterion variables in tests of concurrent or predictive validity. Therefore, it is not clear that the theory linking these to the construct is thought to be sufficiently well developed to term this a test of nomological validity. At best, this must be regarded as very weak evidence for nomological validity--far too weak to support the strong conclusions reached by the authors. With respect to the evidence, no comparison with the more parsimonious one factor model is offered. The question as to how much is being added by going to a two component model is worth addressing.

#### Attitude Construct Validity: Some Issues

The preceding discussion of construct validity and the authors' focus on the attitude construct raised several substantive and methodological issues which it might be helpful to explore.

#### Developing a Nomological Network

Objections were raised earlier to Bagozzi and Burnkrant's selection of scaled behavioral intention and behavior data to test theoretical predictions involving the attitude construct. Since this is a necessary step in establishing construct validity, what sorts of variables might be used? While attitude theory is not particularly well developed, it need not remain so. We might begin with a reexamination of Allport's classic definition of an attitude.<sup>2</sup> An attitude was conceived to be

<sup>2</sup>See McGuire (1969) for a related discussion and many helpful references.

a mental and neural state of readiness to respond, organized through experience and exerting a directive and/or dynamic influence on behavior.

If attitudes are a neural state, physiological variables (e.g., GSR, heart rate) might be used as observables to distinguish people who obtain high vs. low scores on proposed attitude measures, assuming of course that the attitude object used in the study is sufficiently arousing. Those holding rather extreme attitudes (e.g., extreme dislike of a minority group) should respond more intensely to a salience-increasing stimulus. Readiness to respond might, thus, be translated into response latency or intensity (Lott and Lott, 1968; Weiss, 1968). To the extent an attitude represents an organized cognitive structure, certain inference processes (i.e., based on evaluative consistency) may accompany favorable or unfavorable attitudes toward a person or object. Relationships among particular attitudes or attitudes and values might also be predictable on theoretical grounds. One might expect an attitude to exert a directive influence on certain perceptual processes (redefinition of a stimulus, distortion, encoding) as well as on the response side, which, at a minimum, could be measured by something other than the standard pencil and paper set of cognitive response formats (unobtrusive measures, choice behavior, multidimensional scaling).

This is, of course, a much abbreviated list. It is intended to be suggestive of a set of theory-driven relationships which, at present, await more comprehensive conceptualization of the nomological network in which the construct "attitude" is embedded. It is acknowledged that the theoretical links to any one variable are now fairly weak. What is probably needed, therefore, is far greater attention to theory development followed by an investigation of a set of relationships examining various aspects of the nomological network. At the very least, such an approach would guarantee much greater independence among measurement methods as suggested by the Campbell and Fiske paradigm.

#### Levels of Analysis

As discussed earlier, Bagozzi and Burnkrant proposed a two component attitude construct made up of cognitive and affective dimensions. They outlined a theoretical system in which these two dimensions impact directly on behavioral predispositions which, in turn, lead to overt behavior. From their discussion, it seems clear that the term "components" refers to what the authors believed to be a necessary partitioning of the attitude construct and not to a deterministic relationship spanning different psychological levels of analysis.

In discussing Fishbein and Ajzen's version of an expectancy-value model, John and Reve state that attitude "consists of" two components (these are mislabeled in the paper, but that is not crucial to this point), and then it is said that these components "determine one's attitude." Consisting of and determining are not the same thing. This highlights the importance of specifying the level of analysis at which the construct is being defined or explained and not attempting to choose among levels in making construct validity inferences. Cook and Campbell (1976) make the point this way in discussing construct validity in reference to "threats" to the proper labeling of cause and effect in experiments. Such "threats" produce "confounding" in the sense that, "cause and effect can be construed in terms of more than one construct, all of which are stated at the same level of reduction...The reference to the level of reduction is important because it is always possible to 'translate' sociological terms into psychological terms, or psychological terms into biological terms" (1976, p. 238).

Similarly, it is possible to "translate" the construct "attitude" into lower-level constructs which are theor-

ized to be its building blocks. Fishbein and Ajzen define attitude as "a learned predisposition to respond in a consistently favorable or unfavorable manner with respect to a given object" (1975, p. 6). Consistent with this definition Fishbein and Ajzen suggest that attitude "should be measured by a procedure which locates the subject on a bipolar affective or evaluative dimension..." (1975, p. 11). They add, "Beliefs are the fundamental building blocks in our conceptual structure... a person's attitude toward an object is based on his salient beliefs about that object" (1975, p. 14). Accordingly, they describe the relationship between the set of lower-level beliefs and attitude in terms of an expectancy-value model which specifies both the types of beliefs (i.e., beliefs about the object's association with attributes and consequences and beliefs about the evaluation of the attributes and consequences) and the functional relationship among such beliefs.

It is, of course, meaningful to evaluate attitude measures which vary by level of reduction in terms of criterion-related validity. Different levels of measurement may be appropriate depending upon the criterion variable and the purpose of the investigation (e.g., prediction, diagnosticity).

#### Parsimony vs. Completeness

Parsimony has long been regarded as a virtue in theory building; unnecessary constructs and too cumbersome theoretical structures to be replaced by simpler formulations whenever possible. One must wonder, therefore, whether a multicomponent model of attitudes that fits ever so slightly better is to be preferred over a simpler model. Following his extensive analysis of the theoretical and empirical literature dealing with attitudes, McGuire concluded: "Given the less than perfect state of our measuring procedures, the three components have proven to be so highly intercorrelated that theorists who insist on distinguishing them should bear the burden of proving that the distinction is worthwhile" (1969, p. 157). Despite the development of more finely tuned statistical methods for partitioning matrices and accounting for variance, I wonder if that conclusion is any less valid today. The key word may be "worthwhile".

I differentiate in this section between constructs and theory, on the one hand, and applied research on the other. The goals need not be the same. In the latter case, completeness (in the sense of measuring all the variables one believes will prove useful) is much to be valued. It may be no coincidence that many of the social psychologists who are identified with the multicomponent view of attitude were keenly interested in topical issues such as prejudice, attitudes toward the war, political ideology, strategies of persuasion and the like. In seeking to adequately describe the phenomena of interest, it wasn't enough to develop a unidimensional measure of affect, and more elaborate building block models (e.g., information processing, expectancy-value approaches) were still some years away. As a result, not only did researchers (often using survey-type questionnaires) want to know what people knew about the issue, how they felt about it and what types of action they were prepared to take, but also people's intensity of feeling, interconnectedness among beliefs, differentiation of beliefs, relationship to central values, etc. In following this approach a key question is, "What is worth measuring?" The answer is normally decided on criterion-related rather than theoretical grounds.

It may be instructive to contrast the above orientation--which leads to a particularly rich and complete treatment of a construct--with an exceedingly parsimonious approach. Robert Wyer sees an attitude as simply another belief, with no fundamental difference between one belief or another: "A subject's reported

attitude toward an object is interpretable in terms of his judgment of the object's membership in a cognitive category" (1974, p. 24). In other words, having an attitude that bank robbers are "bad" is nothing more than a relationship between membership in the category "bank robbers" and membership in the category "bad". Wyer is expressing a preference for laws of behavior that avoid an Aristotelian emphasis on surface characteristics in favor of a Galileian focus on underlying cognitive process. This emphasis is worth thinking about in approaching the subject of construct validity.

#### REFERENCES

- \* Construct validity and related philosophy of science issues defy easy resolution. When positions are taken they tend to reflect one's orientation to and faith in science. I wish to acknowledge stimulating discussions on these topics with a number of my colleagues at Florida, particularly Gordon Bechtel and Barry Schlenker.
- Standards For Educational and Psychological Tests, (1974), American Psychological Association, p. 29, Washington, D.C.
- D.T. Campbell and D.W. Fiske, "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix," Psychological Bulletin, 56 (March 1959), 81-105.
- T.D. Cook and D.T. Campbell, "The Design and Conduct of Quasi-Experiments and True Experiments in Field Settings," in M.D. Dunnette (ed.), Handbook of Industrial and Organizational Psychology (Chicago: Rand McNally, 1976), 223-326.
- L.J. Cronbach and P.E. Meehl, "Construct Validity in Psychological Tests," Psychological Bulletin, 52 (July 1955), 281-302.
- M. Fishbein and I. Ajzen, Belief, Attitude Intention and Behavior: An Introduction to Theory and Research, Reading, Mass. Addison, Wesley (1975).
- A.J. Lott and B.E. Lott, "A Learning Theory Approach to Interpersonal Attitudes," in A.G. Greenwald, T.C. Brock and T.M. Ostrom (eds.) Psychological Foundations of Attitudes, (New York: Academic Press, 1968).
- W.J. McGuire, "The Nature of Attitudes and Attitude Change," in G. Lindzey and E. Aronson (eds.) The Handbook of Social Psychology, (Reading, Mass., Addison Wesley, 1969).
- R.F. Weiss, "An Extension of Hullian Learning Theory to Persuasive Communication," in A.G. Greenwald, T.C. Brock and T.M. Ostrom (eds.) Psychological Foundations of Attitudes, (New York: Academic Press, 1968).
- R.S. Wyer, Jr., Cognitive Organization and Change: An Information Processing Approach, (New York: John Wiley, 1974).

STRUCTURAL MODELLING: AN APPLICATION FOR TESTING ATTITUDE MODELS AND CONVERGENT VALIDITY

Richard Staelin, Carnegie-Mellon University\*

Abstract

The common theme of this session is validity. Both the Bagozzi-Burnkrant and John-Reve papers are concerned with estimating convergent validity, while the Cattin paper deals with the problem of predictive validity. My discussion concentrates on the first topic, i.e., convergent validity, although the discussion is done within the broader framework of testing and/or validating a theory which is stated in terms of unobserved constructs.

Introduction

A major methodological problem in consumer research is the testing and/or validation of a theory where the theory of interest is stated in terms of unobservable constructs while the variables used to test the model are fallible measures of these constructs. Until recently most analyses of consumer behavior models have not explicitly taken into account this difference between the theoretical constructs and the observed measures. However, in the last few years a new methodology, often referred to as structural equation modelling, has developed which allows consumer researchers to incorporate into their analyses both aspects, i.e., that the theory is in terms of theoretical constructs and the data used to test the theory are composed of fallible measures.

This structural equation modelling approach forms the foundation for the analyses performed by Bagozzi and Burnkrant (B-B) and John and Reve (J-R). It is my personal opinion that this approach has significant potential in testing specific consumer theories. However, as with most "sophisticated" methodological approaches, it is necessary for the user to understand the assumptions required by this approach. One of the goals of this paper is to critically review the assumptions (both implicit and explicit) made by the two sets of authors and to suggest an alternative formulation which circumvents one of the more restrictive assumptions.

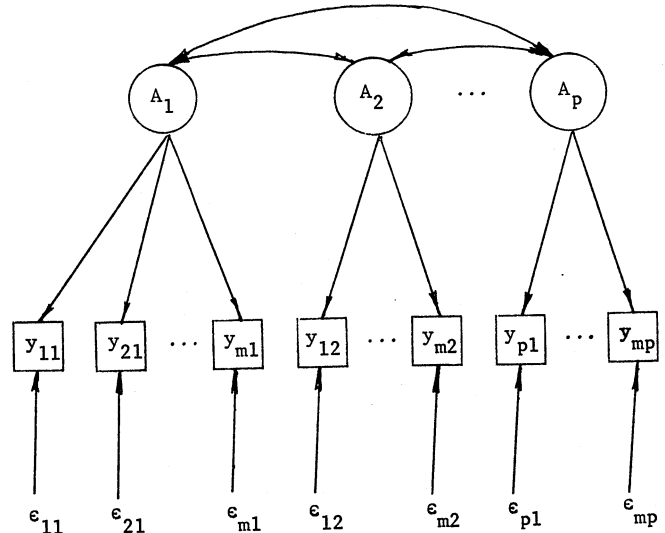
Structural Equation Modelling

Both B-B and J-R are concerned with modelling the concept attitude and the process by which attitude is measured. Both sets of authors start out with the assumption that attitude is a multi-dimensional construct and that the measures used to tap each component are fallible and are correlated either because they measure correlated constructs or because they are obtained via similar data collection methods. Using this general conceptualization both studies rely on previously collected data to estimate the dimensionality of attitude and the validity of the specific measures used to measure each component of attitude.

This conceptualization is represented schematically in Figure 1 where the underlying attitude constructs are designated by circles, the squares represent the fallible measures, the double-headed arrows the links within the individual constructs, the single-headed arrows the postulated causal relationships between the constructs and fallible measures and the  $\epsilon$ 's, the measurement errors associated with the measurement process. The causal model is further specified by assuming a functional form between the constructs and the observed

FIGURE 1

SCHEMATIC REPRESENTATION OF ATTITUDE MODEL AND THE MEASUREMENT PROCESS ASSOCIATED WITH TAPPING EACH COMPONENT OF ATTITUDE



variables. Both sets of authors make the "standard" assumption that this form is linear and additive with the errors assumed to come from a multivariate normal distribution with mean vector equal to zero and a specified variance-covariance matrix. For example, the functional form (e.g., the structural equations) assumed by B-B is

$$y_{ijk} = \lambda_{jk} A_{ij} + \epsilon_{ijk} \quad i = 1, 2, \dots, n; j = 1, \dots, p; k = 1, \dots, m; \quad (1)$$

where  $y_{ijk}$  is the  $k^{\text{th}}$  observed measure for person  $i$  on construct  $j$ ,  $\lambda_{jk}$  is the coefficient of construct  $j$  for measure  $k$ ,  $\lambda_{jk}$  indicating the scaling between the unobserved construct  $A_{ij}$  and the observed measure, and  $\epsilon_{ijk}$  is the measurement error.

The general approach used by both sets of authors to estimate parameters of a model of this form is maximum likelihood. Conceptually this is done by trying to find that set of parameters which best reproduces the covariance (or correlation) matrix of the observed data. To get a feel of how this is done, let us address our attention to a few elements in the covariance matrix of the  $y$ 's assuming that the measurement process is modelled as stated in Equation 1. For example, the variance of the first observed measure for construct 1, i.e.,  $\text{Var}(y_{11})$  can be expressed in terms of the model's parameters as follows:

$$\text{Var}(y_{11}) = \lambda_{11}^2 \text{Var}(A_1) + \text{Var}(\epsilon_{11}) + 2 \lambda_{11} \text{Cov}(A_1 \epsilon_{11}). \quad (2)$$

Likewise the covariance between  $y_{11}$  and the first measure of the second construct  $y_{21}$  is

$$\begin{aligned} \text{Cov}(y_{11}, y_{21}) &= \lambda_{11} \lambda_{21} \text{Cov}(A_1 A_2) + \text{Cov}(\epsilon_{11} \epsilon_{21}) \\ &+ \lambda_{11} \text{Cov}(A_1 \epsilon_{21}) + \lambda_{21} \text{Cov}(A_2 \epsilon_{11}), \end{aligned} \quad (3)$$

while the covariance between  $y_{11}$  and the second measure of the first construct  $y_{12}$  is

$$\begin{aligned} \text{Cov}(y_{11}, y_{12}) &= \lambda_{11} \lambda_{12} \text{Var}(A_1) + \text{Cov}(\epsilon_{11} \epsilon_{12}) \\ &+ \lambda_{11} \text{Cov}(A_1 \epsilon_{12}) + \lambda_{12} \text{Cov}(A_1 \epsilon_{11}). \end{aligned} \quad (4)$$

In an analogous manner each element of the observed variable covariance matrix can be represented in terms of the unobserved model parameters. It is often the case in analyses of this type that the number of possible model parameters obtained in the above fashion exceeds the number of observations. Consequently, it is not possible to uniquely identify the parameter without limiting the parameter space by making some assumptions about the permissible values for certain of these parameters. For example, the assumptions used in the B-B paper are as follows.

1. The variances of the unobserved constructs (e.g.,  $\text{Var}(A_i)$ ) are equal to one. This assumption is normally costless since the constructs are never measured and thus can be arbitrarily rescaled to yield any variance desired.
2. The covariances between the error terms and the constructs (e.g.,  $\text{Cov}(A_i \epsilon_{11})$ ) are equal to zero. This assumption is not costless. One might postulate conditions where the errors associated with the measurement process are a function of the level of the construct. For example, people with a high (favorable) attitude toward a product might try to please the interviewer and thus tend to give an answer which is more favorable than their true attitude while the converse would hold for respondents who have an unfavorable opinion about the product.
3. The covariance between one construct and an error term associated with another construct (e.g.,  $\text{Cov}(A_1 \epsilon_{21})$ ) is equal to zero. This does not seem to be too restrictive. It is hard to imagine situations where the error term associated with one construct would be correlated with another construct.
4. The covariance between two error terms (e.g.,  $\epsilon_{11}$  and  $\epsilon_{21}$ ) are equal to zero. This is probably the least palatable of the assumptions made by B-B since it is easy to conjure up situations where the error terms for two measures are correlated because they either are designed to tap the same construct or they are obtained by a similar scaling method. The analysis of J-R is aimed at circumventing this problem by postulating a methods factor.<sup>1</sup> Another approach compatible with structural equation modelling postulates error factors which are correlated unobserved constructs (Joreskog and Sorbom, 1977). I use this latter approach in developing an alternative model to that proposed by J-R.

Applying the above four assumptions to Equations 2-4 yields equations of the following form:

<sup>1</sup> It is interesting to note that the model used by J-R was previously proposed by Bagozzi (1978), although he does not use this model in the current B-B paper.

$$\text{Var}(y_{11}) = \lambda_{11}^2 + \text{Var}(\epsilon_{11}), \quad (5)$$

$$\text{Cov}(y_{11}, y_{21}) = \lambda_{11} \lambda_{21} \text{Cov}(A_1 A_2), \quad (6)$$

$$\text{Cov}(y_{11}, y_{12}) = \lambda_{11} \lambda_{12}. \quad (7)$$

It should be noted that these four assumptions are not the only restrictions placed on the model's parameters. Instead of using the covariance matrix, B-B use the correlation matrix for the  $y$ 's. This means that the variance of each  $y$  is one. As a consequence, the authors are implicitly imposing another set of restrictions on the parameters of the structural equations (i.e.,  $\lambda_{ii}^2 + \text{Var}(\epsilon_{ii}) = 1$ ). In other words the parameter space is further restricted by requiring  $\lambda_{ii}$  to be dependent on the value of the variance of  $\epsilon_{ii}$ . Although this assumption is not necessarily restrictive,<sup>2</sup> it is important to recognize this implicit assumption when determining a) the number of free parameters estimated and b) the degrees of freedom associated with any statistical tests of the model.

The next step in the analysis is to assume a specific distributional form for the  $y$ 's conditional on the unobserved constructs (both sets of authors assume this distribution to be the multivariate normal) and then find that set of parameter values (e.g., the  $\lambda$ 's,  $\text{Var}(\epsilon_{jk})$ 's and the  $\text{Cov}(A_i A_j)$ 's) which maximize the probability of getting the observed covariance matrix. These parameter estimates are often obtained via a computer program called LISREL (Joreskog and van Thilo, 1972), which also yields statistics that can be used to test the null hypothesis that the postulated model is correct.<sup>3</sup> It is via this computer program or a variant of this program that both B-B and J-R a) estimated the model parameters b) tested for the convergent validity of specific attitude measures and c) tested for the veracity of their multiple component model of attitude.

Although it should go without saying, it should be noted that the above methodology can not be used to prove a particular theory is correct. As is true with any hypothesis testing, statistical tests of hypotheses are capable only of rejecting a given theory (subject to a type-I error) or conversely saying that the observed data is consistent with the postulated theory. Thus even though the final models postulated by B-B and J-R are not rejected, it is not proper to say that these models are correct. In other words, B-B and J-R do not show that the measures used have convergent validity or that attitude is two dimensional, they only show that they can not reject these statements.

In a somewhat similar vein it must be recognized that the structural modelling methodology does not prove that the model as postulated is valid (i.e., does not prove construct validity). However, the approach is consistent with the philosophy of science notion of testing a theory in that a) the structural equations explicitly state the theory to be tested and b) the methodology allows the theory to be rejected.

<sup>2</sup> I later show a case where this assumption is restrictive.

<sup>3</sup> Specifically, the null hypothesis is that the true model generating the data is the postulated model and that parameters of the model are equal to values estimated via the maximum likelihood estimation. The alternative hypothesis is that the null hypothesis is not true.

## Results

The B-B and J-R papers postulated a number of structural models of attitude and the measurement process. The simplest of these models is stated in Figure 1 of B-B. Using a data base collected by Fishbein and Ajzen (1974) they estimated the parameters of the model using LISREL. The parameter estimates and relevant test statistics for this model are given in Table 1 of their paper. They conclude, based on the  $\chi^2$  value for the two attitude factor model, that this model can not be rejected and thus "convergent validity is established for the two-factor affective/cognitive model of attitudes." As previously mentioned, this is too strong a statement since the  $\chi^2$  test only indicated that the data were consistent with their model.

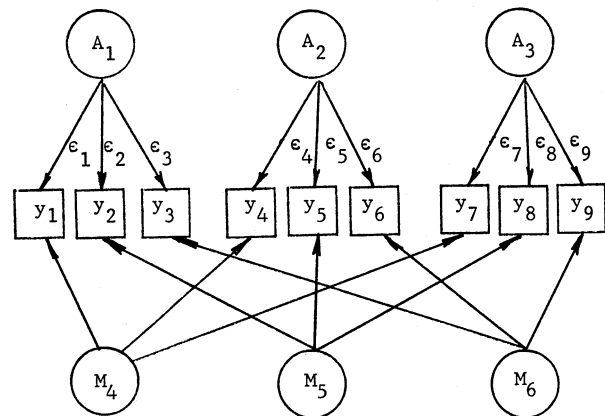
On a more general level it is interesting to look at the estimated correlation between the two components of attitude ( $\phi_{A_1 A_2}$  in their notation). This correlation is estimated as being .824 using data from the self-reported behavior sample and .927 from the behavioral intention sample. In either case, the correlation is high indicating that although there may be two components to attitude toward religion, these two components are highly correlated.

It is possible to test whether the two components of attitude can be lumped into one construct without any loss of statistical significance. Joreskog and Sorbom (1977) show that asymptotically, the difference between  $\chi^2$  value<sup>4</sup> associated with a particular model and the  $\chi^2$  associated with a model which is a subset of that model, is also distributed  $\chi^2$  with degrees of freedom equal to the difference in the number of parameters in the two models. Since B-B report the  $\chi^2$  values for both the one factor (construct) model and the more general two construct model, it is possible to test whether the two highly correlated components can be combined into one scalar quantity.<sup>5</sup> The  $\chi^2$  values (which have one degree of freedom) are 21.61 (= 23.91 - 2.30) and 5.58 (= 13.98 - 8.30) respectively for the two different samples. In the first case the  $\chi^2$  value is significant, indicating the hypothesis that the true model consists of just one component of attitude should be rejected. However, for the second sample, the  $\chi^2$  value is not significant at the .10 level. Thus, contrary to the conclusions of B-B, the results from the two samples do not "establish" a two factor model, but instead yield mixed interpretations.

The approach of J-R was similar to B-B in that they also postulated a model which stated that attitude was a multi-dimensional concept and that the measures used were fallible. However, their model of the measurement process reflected the fact that the observed responses could also be affected by the method used to obtain the responses. More specifically, they postulated that the observed measures  $y_{ijk}$  are influenced not only by the level of the particular construct ( $A_{ij}$ ) they were intended to measure, but also that they were influenced by the method ( $M_{ik}$ ) used to obtain the measure. This conceptualization is schematically displayed in Figure 2 where the three A's represent three

FIGURE 2

SCHEMATIC REPRESENTATION OF THE JOINT INFLUENCES OF METHOD AND ATTITUDE ON OBSERVED RESPONSES



components of attitude and the three M's the three different methods used to obtain the nine y responses. Thus (using their notation) the observed response for person  $i$  for measure  $l$  on trait  $j$  using method  $k$ , is

$$y_{il} = \lambda_{lj} A_{ij} + \lambda_{lk} M_{ik} + \epsilon_{il}, \quad \begin{aligned} l &= 1, 2, \dots, 9, \\ j &= 1, 2, 3 \\ k &= 4, 5, 6 \\ i &= 1, 2, \dots, n, \end{aligned} \quad (8)$$

where the covariance matrix of the unobserved constructs (these being the A's and the M's) are allowed to be correlated.

The model as stated in Equation 8 has some interesting properties. First, the term  $\lambda_{lk} M_{ik}$  "corrects" for method bias by postulating that "everyone has an unobserved (but existing) true score on a method which is independent of any attitude measure. I know of no theoretical reason why a respondent would have such a score, nor do the authors provide the reader with one. In other words, although there are no theoretical grounds for invalidating their model of reality, I do not find the model intuitively appealing. Second, the authors assume that the effect of this "measurement bias" is linear and additive with the attitude score. Again, there is no justification for such an assumption.<sup>6</sup>

The second major observation about Equation 8 is more technical, but is probably more important. Although I haven't been able to prove it, I don't think their general model (i.e., Model 1 of their paper) is identified. This means that the estimates they report are not unique. My general approach to determine if the model was identified was to write out the covariance matrix of the  $y$ 's in terms of the model parameters (in a manner similar to my Equations 2 through 4). Then after reducing the parameter space by making the assumptions similar to those

<sup>4</sup>The  $\chi^2$  value is obtained from  $-2 \ln \lambda$  where  $\lambda$  is the ratio of the likelihood function of the model being estimated to the likelihood function of the model where the parameter values are not restricted.

<sup>5</sup>B-B did not perform this type of test. Instead they looked at the  $\chi^2$  value in each sample associated with the null hypothesis that attitude is composed of one component. In both cases, this  $\chi^2$  value allowed them to reject the null hypothesis, i.e., they concluded the one factor model of attitude was incorrect.

<sup>6</sup>In fairness to the authors, I must acknowledge that most model builders make the assumption of linear and additive terms since in general this assumption seems to be very robust.

made by B-B, I attempted to uniquely determine the parameters in terms of the observed y's. To the best of my ability I was not able to do this. Moreover, I don't think it is possible to uniquely identify all of the parameters. I suspect that the authors did not attempt this analysis either but instead assumed their model was identified simply because the number of parameters estimated was less than the number of observations.

The problem of identification is always difficult in models as complex as that postulated by J-R. One of the standard ways of handling the identification problem is to finesse the type of analysis I attempted and instead input the specified model into the LISREL program and let the program "determine" identification via the information matrix. For identified models the information matrix will almost certainly be positive definite. Conversely, if the information matrix is not positive definite, the specified model is almost certainly not identified (Joreskog and Sorbom, 1977). Since positive definite matrices can be inverted while non positive definite matrices can not, inversion of the information matrix implies that the information matrix is positive definite. Thus, inversion is taken as "proof" that the model is identified. Unfortunately, inversion of the information matrix is only a necessary condition (in contrast to a sufficient condition) for identification. Thus even though I understand why J-R might have resorted to the above "rule of thumb," this rule may have misled them into believing that their general model is identified.

As further "proof" that this model is not identified, note that when the authors further restricted the general model (yielding their Models 2, 3 and 4) the program failed to converge. Lack of convergence is normally a sign that the model is not identified. Since it is impossible for a subset of a model not to be identified when a more general model is identified, my priors are that the general model is also not identified. Thus, their Model 1 results displayed in Figure 3 of their paper are meaningless.

The net result of the above discussion is that I believe it is impossible to identify (solve) for a general three-trait, three-method model where the methods and attitudes are allowed to be correlated using the J-R approach. However, I was able to develop a model which is identified (in fact many of the parameters are over-identified) that assumes that the observed score is affected by both the respondent's attitude and the type of measurement method used. This model is presented next.

#### Alternative Approach to J-R Model

Assume the data of Ostrom (1969) (i.e., the data used by J-R) where three measures are obtained for each of three different attitude components. Let  $y_{ijk}$  represent the observed measure for person  $i$  on attitude component  $j$  using measure  $k$ . Let  $A_{ij}$  be person  $i$ 's attitude score on component  $j$  and let  $M_{ik}$  be person  $i$ 's measurement score on measure  $k$ . Next, postulate the model

$$y_{ijk} = A_{ij} + M_{ik}, \quad \begin{aligned} l &= 1, 2, 3, \\ j &= 1, 2, 3, \\ k &= (l-1)*3 + j, \end{aligned} \quad (9)$$

where  $A_1, A_2$  and  $A_3$  are allowed to covary as are the sets  $\{M_1, M_2, \text{ and } M_3\}$ ,  $\{M_4, M_5 \text{ and } M_6\}$  and  $\{M_7, M_8 \text{ and } M_9\}$ ; otherwise, the constructs are assumed to be independent. In words, Equation 9 says that a person's score is made up of two random variables, an attitude score, and a "measurement score" which represents the measurement error. A person's score on any one

attitude is correlated with the score on the other two attitudes but is independent of the measurement error. Similarly, the measurement error is correlated with errors on other measurements using the same method but independent of those not using the same method. Measures 1, 2 and 3 represent the Thurstone method, 4-6 the Likert method and 7-9 the Guttman method. The covariance matrix of the y's for this model is given in Table 1 where  $\varphi_{ij}$  is the covariance between attitude components  $i$  and  $j$  and  $\psi_{ij}$  is the covariance between measurement errors  $i$  and  $j$ . Clearly, the  $\varphi_{ij}$ 's are over-identified and the  $\psi_{ij}$ 's are just identified.<sup>7</sup>

The above model has some nice properties. First, it is possible to test for the dimensionality of the unobserved construct attitude. Second, it acknowledges that the observed measures may be correlated for two reasons, i.e., the measures tap correlated constructs and they are derived from similar methods. Finally, it partitions the response variance of  $y_{jk}$  into two components, i.e.,  $\varphi_{jj}$  and  $\psi_{kk}$ .

As with previously discussed models, this specification makes a number of assumptions. First, it assumes that the person's measurement error is not affected by the level of a person's attitude. Thus, for example, a person with a favorable attitude is no more likely to have a high (low) measurement error than one with an unfavorable attitude. Second, it assumes that measurement errors associated with different methods are independent. Finally, it assumes that the attitude score and measurement error are additive. Although I offer no proof as to the veracity of these assumptions they do not seem to strain one's credibility.

It should be noted that the model is stated in terms of covariance versus correlations. Moreover, it is not possible to arbitrarily rescale the model to get the y's to have unit variances without imposing restrictions on the magnitude of the variances of the unobserved variables. Thus, if the correlation matrix were used instead of the covariance matrix, the parameter space would be further restricted so that

$$\varphi_{11} + \psi_{11} = 1, \quad \varphi_{11} + \psi_{44} = 1, \quad \text{and} \quad \varphi_{11} + \psi_{77} = 1$$

$$\text{or} \quad \psi_{11} = \psi_{44} = \psi_{77}.$$

$$\text{Likewise} \quad \psi_{22} = \psi_{55} = \psi_{88} \quad \text{and} \quad \psi_{33} = \psi_{66} = \psi_{99}.$$

Clearly this is not a desired model property since there is no reason to believe that all the measurement errors using the same method have the same variance.

The above indicates that it is improper to use the correlation matrix as input. Unfortunately, Ostrom reported only the correlation matrix. Although there is some indication that he tried to insure that the variances for the observed scores for a given method are equal, use of the correlation matrix to estimate model parameters is incorrect. However, given the alternatives of doing nothing or doing something, I elected for the latter since it appeared that the variances of the y's for a given method were approximately equal. Consequently, I did not feel that there was too much error in this case in using the correlation matrix to estimate

<sup>7</sup>To see this, note that, for example, it is possible to uniquely estimate  $\varphi_{11}$  from three different elements in the matrix, i.e., row 4, column 1; row 7, column 1; and row 7, column 4. Once the  $\varphi_{ij}$ 's are estimated, the 18  $\psi_{ij}$ 's can be estimated from the 18 different entries  $\psi_{ij}$  which include these parameters.

TABLE 1

## COVARIANCE MATRIX IN TERMS OF MODEL PARAMETERS

$\varphi_{11} + \psi_{11}$									
$\varphi_{12} + \psi_{12}$	$\varphi_{22} + \psi_{22}$								
$\varphi_{13} + \psi_{13}$	$\varphi_{23} + \psi_{23}$	$\varphi_{33} + \psi_{33}$							
$\varphi_{11}$	$\varphi_{12}$	$\varphi_{13}$	$\varphi_{11} + \psi_{44}$						
$\varphi_{12}$	$\varphi_{22}$	$\varphi_{23}$	$\varphi_{12} + \psi_{45}$	$\varphi_{22} + \psi_{55}$					
$\varphi_{13}$	$\varphi_{23}$	$\varphi_{33}$	$\varphi_{13} + \psi_{46}$	$\varphi_{23} + \psi_{56}$	$\varphi_{33} + \psi_{66}$				
$\varphi_{11}$	$\varphi_{12}$	$\varphi_{13}$	$\varphi_{11}$	$\varphi_{12}$	$\varphi_{13}$	$\varphi_{11} + \psi_{77}$			
$\varphi_{12}$	$\varphi_{22}$	$\varphi_{23}$	$\varphi_{12}$	$\varphi_{22}$	$\varphi_{23}$	$\varphi_{12} + \psi_{78}$	$\varphi_{22} + \psi_{98}$		
$\varphi_{13}$	$\varphi_{23}$	$\varphi_{33}$	$\varphi_{13}$	$\varphi_{23}$	$\varphi_{33}$	$\varphi_{13} + \psi_{79}$	$\varphi_{23} + \psi_{89}$	$\varphi_{33} + \psi_{99}$	

the model parameters. These estimated model parameter values are reported in Table 2.

The results indicate that it is not possible to reject the hypothesis that three components of attitude exist and that measurement process is correct as modelled. The estimated correlations between the three attitude components indicate a high degree of commonality between the element (i.e., between .79 and .91), although the estimated correlations are lower than those reported by J-R. In general, the Likert method seemed to produce the smallest error variance, with Thurstone being next best, followed by the Guttman method. Interestingly, the correlation between the scores on measures using the same method were usually very small. The only method to exhibit significant (managerial and statistical) correlations was the Likert method.

#### Summary of Convergent Validity Papers

The above discussion is intended to supplement B-B's and J-R's presentation of the structural equation approach used by both sets of authors to test for convergent and discriminant validity. One of the major themes of this discussion is that if researchers are to use this approach they must first carefully think out all the assumptions which are implicitly or explicitly tied to the technique. In my opinion the J-R paper does not reflect this careful thinking in that they present a model which has no theoretical justification nor is it identified.

A second aim of my discussion is to put forward a model which simultaneously takes into account the influences on the observed responses of the constructs being tested, any relationships between the constructs and the methods used to obtain these measures. This structural model is an outgrowth of a model suggested by Bagozzi (1978) and implemented by J-R. Unfortunately the B-B paper does not use either the model proposed by Bagozzi or a variation of this model when testing the validity of the two component models of attitude even though the measures used to test the theory were obtained using the same method (see Figure 3 of B-B paper). Consequently, it is not possible for them to test if the estimated relationship between constructs

A<sub>2</sub>(affect) and B (behavioral intentions) is due to the measurement technique or the basic underlying relationship between the two variables. Hopefully, future work in this area will control for such confounding influences by incorporating my approach or some improved variation.

#### Predictive Validity

The emphasis of the Cattin paper is on obtaining a model which has high predictive power. This is in contrast with the two previously discussed papers which were conceived with testing the validity of measures used to test a theory. Thus the only real commonality between Cattin's paper and the papers of B-B and J-R is the word "validity."

The Cattin paper first reviewed six formulae which estimate the correlation between an observation and the prediction of that observation (based on the regression model). It then presented an example which indicates for the data set selected that it is better to use an estimate of the population cross validation correlation than the more commonly used sample correlation estimate when selecting between regression models. Cattin offered no proof as to the generalizability of these results although he did review a series of Monte Carlo studies aimed at determining the properties of each estimator.

My major concern with the Cattin paper is the emphasis on predictive power versus theory as the criterion for selecting a model. Thus I don't believe a model normally should be selected just because it has a high population cross validation correlation. Instead, in most consumer behavior research, a researcher should start with a theory. This theory will determine the predictor variables and the functional form of the model. If the researcher wants to conduct an F test to determine if a subset of the model adequately represents the data, this is certainly permissible. However, selecting a model just because it has the highest predictive power is atheoretical and should be avoided in most consumer research settings. Thus, I am not in disagreement with Cattin over any technical issue. Instead, I question the value of the cross validation correlation

coefficient in most consumer research situations. The only time this statistic would be of value is when a researcher wanted to develop a model that predicts well for different data sets. Even then, I would not just use the cross validation correlation coefficient in selecting the model, but also blend in my opinions as to the soundness of the alternative models.

References

\* I wish to acknowledge the helpful comments of Robert Avery, V. Joseph Hotz, Andrew Mitchell and Robert Redinger.

R. P. Bagozzi, "The Construct Validity of the Affective, Behavioral and Cognitive Components of Attitude by Analysis of Covariance Structures," Multivariate Behavioral Research, 13 (January, 1978), 9-31.

M. Fishbein and I. Ajzen, "Attitude Towards Objects as Predictors of Single and Multiple Behavioral Criteria," Psychological Review, 81, (1974), 59-74.

K. G. Joreskog and D. Sorbom, "Statistical Models and Methods for Analysis of Longitudinal Data." In Latent Variables in Socioeconomic Models, ed. by D. J. Aigner and A. S. Goldberger, North Holland, Amsterdam 1977.

K. G. Joreskog and M. van Thillo, LISREL: A general computer program for estimating a linear structural equation system involving multiple indicators of unmeasured variables. Educational Testing Service Research Bulletin, (1972), 72-156.

T. M. Ostrom, "The Relationship Between the Affective, Behavioral, and Cognitive Components of Attitudes," Journal of Experimental Social Psychology, 5 (1969), 12-30.

TABLE 2

ESTIMATED PARAMETERS FOR OSTROM DATA

VARIABLE	VARIABLE NAME	VARIANCE	CORRELATION
$\Phi_{11}$	Affect	.65	
$\Phi_{22}$	Conative	.66	
$\Phi_{33}$	Cognitive	.67	
$\Phi_{12}$	Affect-Conative		.79
$\Phi_{13}$	Affect-Cognitive		.91
$\Phi_{23}$	Conative-Cognitive		.86
$\Psi_{11}$	Thurstone-Affect	.31	
$\Psi_{22}$	Thurstone-Conative	.43	
$\Psi_{33}$	Thurstone-Cognitive	.34	
$\Psi_{44}$	Likert-Affect	.25	
$\Psi_{55}$	Likert-Conative	.23	
$\Psi_{66}$	Likert-Cognitive	.24	
$\Psi_{77}$	Guttman-Affect	.58	
$\Psi_{88}$	Guttman-Conative	.38	
$\Psi_{99}$	Guttman-Cognitive	.36	
$\Psi_{12}$			.001
$\Psi_{13}$	Thurstone Methods		-.001
$\Psi_{23}$			.016
$\Psi_{45}$			.176
$\Psi_{46}$	Likert Methods		.072
$\Psi_{56}$			.158
$\Psi_{78}$			-.000
$\Psi_{79}$	Guttman Methods		-.001
$\Psi_{89}$			.000

$\chi^2 = 15.06$

d.f. = 12

p = 24



THE EFFECTS OF PRODUCT INVOLVEMENT AND TASK DEFINITION  
ON ANTICIPATED CONSUMER EFFORT

Keith Clarke (student), University of Illinois  
Russell W. Belk, University of Illinois

Abstract

Two major determinants of the fervor and diligence with which a consumer prepares to make a purchase are thought to be the degree of involvement with the item to be purchased and the degree of involvement in the situation in which the purchasing occurs. This paper reports an experiment to assess the manner in which inherent product involvement and situational task importance affect anticipated consumer purchase effort. Both factors were found to independently increase estimations of purchase effort, but the effect of task importance was negligible in the case of high involvement products. This finding is interpreted as a ceiling effect on overall involvement level.

The concept of consumer purchasing effort has received relatively little attention in consumer research. Cardozo and Bramel (1969) treated the expenditure of effort in consumer decision-making as a form of behavioral commitment to the purchase. In this view, greater effort in a consumer purchase decision increases the perceived importance of the purchase decision and thereby increases the potential for post-purchase cognitive dissonance. The dissonance created by expending an amount of effort which is not commensurate with the degree of satisfaction the product is able to provide, may then be reduced through an inflated post-purchase evaluation of the chosen product. Cardozo (1965) and Woodside (1972) found support for this hypothesis in experiments manipulating the amount of effort required in consumers' prescribed information search activities. However, in much of consumers' normal purchasing activities the search task is much less structured and the relationship between decision importance and purchasing effort may be in the opposite direction from the one set up in these studies. That is, rather than greater search effort causing consumers to perceive a purchase as being more important, greater purchase importance should cause consumers to expend more search effort. This was the primary hypothesis motivating the present research.

Product Involvement and Task Involvement  
as Determinants of Decision Importance

Two separate aspects of the importance of a purchase decision are the amount of product involvement and the nature of the task definition involved in the purchase. Product involvement, as used here, is purchase item-specific but not purchase situation-specific.<sup>1</sup> Thus, for a given individual, a set of products and services might be arrayed in terms of their inherent involvement. While there are individual differences in levels of involvement with a given product, with a relatively homogeneous population the rank orders of involvement with an array of products are expected to be reasonably constant. Hupfer and Gardner's (1971) measures of

product involvement among college students support this expectation, with certain products such as automobiles being almost uniformly more involving than other products such as facial tissues. Ratchford and Andreasen (1974) reached similar conclusions in their examination of the perceived importance of decisions to purchase various services by new residents to a community.

Howard and Sheth (1969, p. 73) predicted that greater search effort will be expended for high importance products. Consistent with this hypothesis, Katona and Mueller (1954) found strong differences in consumer decision effort ("deliberation") between products of such apparently differential involvement as sports shirts (low) versus major appliances (high). However, it is not clear that product involvement was the primary difference between these products, and various authors (e.g. Sherif and Cantril, 1947) have suggested that clothing is highly involving. It is therefore necessary to test any operationalization of product involvement and to avoid generalizing to dissimilar populations.

A second determinant of decision importance is task involvement. Belk (1975) suggested that task-defining features of a consumer purchase situation arise from goals relating to information gathering or product selection, and from the usage situations envisioned for products which are relevant to these goals. In other words, the task is defined by the consumer's intentions at a particular time and place. The task may be highly involving either because it entails important immediate goals (e.g., "find a coat which is the least expensive brown wool coat in town"), or because the intended usage situation involves important goals (e.g. "find a dress to wear to the prom"). Although task definitions for purchase situations may also differ in other respects such as their difficulty, specificity, and complexity, the present study is concerned only with task involvement.

Some support for the hypothesized relationship between task importance and search effort is found in a study by Grønhaug (1972). Grønhaug compared the information search activities of recent buyers of tableware who bought the product either as a gift (higher task involvement) or for personal use (lower task involvement). Those who had purchased the tableware as a gift were found to have considered more alternatives, visited more shops, sought more advice from both dealers and others, and studied more brochures. Based on these results and prior studies of the task involvement levels of gift-giving (e.g. Belk, forthcoming), the present study employs the gift/personal use dichotomy to manipulate task involvement.

2 This was a composite index formed by adding measures of perceived circumspectness, information seeking activity, number of features considered, number of brands considered and importance of price.

1 Howard and Sheth (1969) used the term "importance of purchase," Lastovika (1976) used the term "issue involvement," and Rothchild (1977) used the term "enduring involvement" to refer to this type of involvement.

## Involvement, Arousal, and Effort

A key concept integrating product involvement and task involvement as well as linking them to effort expenditure, is arousal. In an approximate sense, arousal may be seen as the level of energy released to deal with a person's present situation. While product and task involvement are not the only determinants of arousal, in a focused activity like product choice and purchasing they are generally dominant. The general expectation then is that with greater levels of arousal created by higher product involvement, task involvement, or both, the amount of effort devoted to information search and processing increases. It should be noted, however, that most theorists dealing with arousal do not postulate a straight linear relationship between level of arousal and level of efficient effort expenditure. For instance, based on the work of Hebb (1966), Berlyne (1967), and others, Hansen (1972) predicts the greatest amount of information search and deliberation at very low and moderately high levels of arousal, and the lowest amount of effort expenditure on these activities at medium levels of arousal. Despite the theoretical and empirical support for the non-linear relationship between arousal and effort expenditure, the series of experiments conducted by Hansen (1972) supports the view that most consumer purchase activities fall within the range of arousal which avoids the extremes at which arousal may no longer be positively associated with effort. Similarly it is assumed that the levels of involvement created by the present study's product and task involvement conditions are such that there is a direct although not necessarily linear relationship between arousal and consumer purchase effort expenditure.

Consumer purchase effort in response to arousal level may be manifested in several different ways. Effort is spent when the consumer internally recalls and processes information as well as when there are more overt signs of information acquisition. As Newman (1977) points out, external information search can also take on a number of forms and appearances. The most frequently used measure of external consumer search is the number of retail stores visited prior to purchase, and this is the one of the measures adopted for the current study. Nearly all measures of external information search involve active deliberate information search, although recent extensions of the time budget diary method by Venkatesan and Arndt (forthcoming) obtain some measures of passive non-deliberate information acquisition. Because the present study focuses on anticipated effort expenditure, its measures concern deliberate information search. A second measure of consumer effort expenditure used in the current study is deliberation time exclusive of travel time. By excluding travel time the measure ignores locational factors idiosyncratic to the data collection city, and provides a measure of largely internal search and processing activity efforts. The price a consumer expects to pay for a product constitutes a third type of measure in this study. Although price is not a reflection of effort expenditure directly, since effort must usually be expended for the consumer to obtain money, spending more money is an indirect expenditure of effort.

## Hypotheses

The three basic measures of effort expenditure just noted are hypothesized to be related to the two types of involvement in the following ways:

1. More stores will be shopped, more time will be spent; and more money will be spent under high task involvement conditions than under low task involvement conditions.

2. More stores will be shopped, more time will be spent, and more money will be spent for high involvement products than for low involvement products.

3. The effects of task involvement and product involvement will be additive with no interaction.

While each of the hypothesized effects in the first two hypotheses have been discussed, there is an alternative to hypothesizing that all three types of effort expenditure will increase with involvement. The alternative is to consider the three types of effort as substitutes, so that an increase in one type of effort can lead to a reduction in the others. For instance, a consumer might rely on a higher priced product as an assurance of satisfaction and thereby reduce the necessity of additional information search. However, considering that the source of consumer arousal is involvement with the product or task, it seems more reasonable that the consumer would be willing to spend more money and more information search effort in order to assure a more satisfactory purchase. Furthermore, in line with the Cardozo (1965) findings, the expenditure of one type of effort may itself increase the overall levels of purchase involvement and arousal which would further increase the other types of effort expenditure. The third hypothesis assumes that regardless of the level of inherent product involvement, giving the product as a gift (high task involvement) increases the overall level of arousal and causes more effort to be expended.

## Method

A two by four factorial design was used. Task involvement was manipulated by informing subjects that the product was to be purchased for personal use or else as a gift for a good friend. Product involvement was operationalized by the use of four products. Two of the products (bubble bath and a blanket) were chosen to represent low involvement and two (jeans and a record album) were chosen to represent high involvement products. These choices were based on a series of direct and indirect measures of involvement with various products conducted by Lastovika (1976) and in pretests conducted by the authors on the same general subject pool as the current study.

In the pretest conducted by the authors, subjects also provided estimates of a low, a medium, and a high level for each of the three dependent measures. That is, subjects in the pretest estimated a low, medium, and high number of stores to visit, a low, medium and high amount of non-travel time to spend in deciding, and a low, medium, and high amount of money to spend for each of the products. These measures were obtained for each product without reference to the task situation. The mean of the "medium amount" was used as the central point in a scale measuring the amount of a type of effort the experimental subjects would be willing to expend in selecting each product. The points on either side of these central points for each scale were selected by taking the median of the high and low amounts for each product. Medians were used so as not to choose points skewed by any extremely high or low values given in the pretest. On each scale, two additional points were added by including the categories "less than (the low amount)" and "higher than (the high amount)". Through the use of these scales it was possible to make comparisons of the objective ("real choice") levels of effort expenditure across the products, even though the base levels of these effort expenditures differed across the products. In an effort to measure subjective ("relative choice") levels of effort expenditures, three additional questions were employed measuring effort expenditure on seven-point scales labeled from "very little" or "very few" to "very much" or "very many".

The same results were expected from measures of both subjective and objective effort expenditures.

Subjects were 56 undergraduate female students. They were randomly assigned to two treatment conditions subject to the constraints that each receive both task conditions but not receive the same product. It was felt that having each subject engage a different task and a different product would minimize the problem of repeated measures effects. To test for this, the data were first analyzed as a two by four by two design with the third factor being whether the responses were to the first product/task or the second product/task. There were no significant main effects or interaction effects involving this third factor so it was felt that there had been

no sensitization problem and the data were collapsed into the two by four design.

Results

The results of the study are summarized in Table 1 and Figure 1. The main effect for task involvement was significant for all six dependent measures, with effort expenditures for gifts reported as higher than expenditures for products for personal use. The second hypothesis was also generally supported, with all relative choice effort measures significant but with numbers of stores the only significant real choice measure. At least on the perceived effort scales then, the amount of effort was greater for the high involvement products.

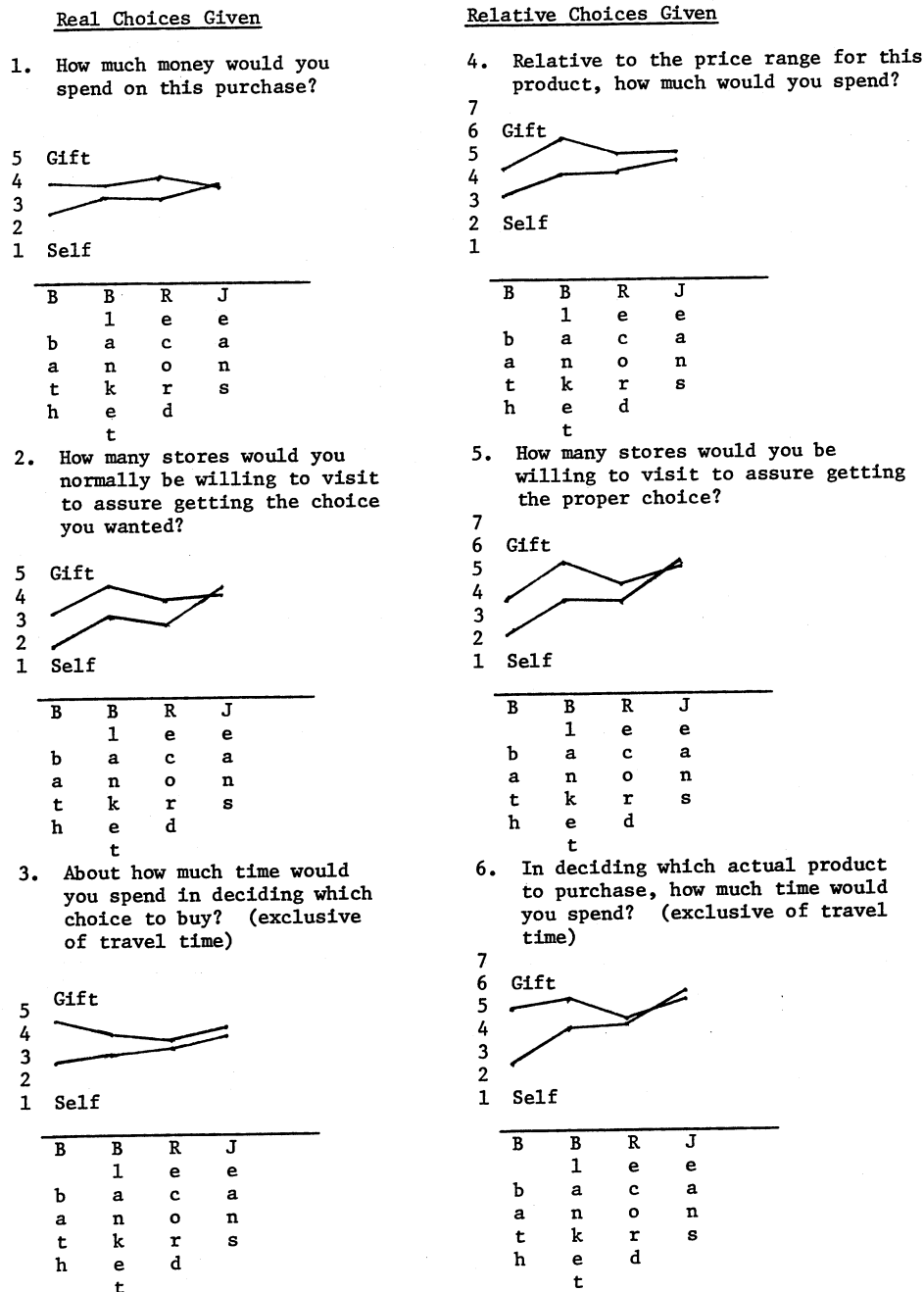
TABLE 1  
ANOVA SUMMARY (F LEVELS)

EFFECT	df	Real Choices			Relative Choices		
		Money	Stores	Time	Money	Stores	Time
Task Involvement	1	26.3	18.0	17.8	27.7	11.6	10.2**
Product Involvement	3	1.4NS	14.2	1.6NS	9.5	12.8	6.1
Interaction	3	4.2**	3.5*	3.7%	1.7NS	2.2NS	3.9**
Error	104						
<hr/>							
Low vs High Involvement products for personal use	1	4.3*	13.2	4.7*	13.0	16.3	14.5
<hr/>							
Low vs High Involvement products as a gift	1	.1NS	.ONS	3.8NS	.6NS	.5NS	.4NS
<hr/>							
Purchase for Self vs Purchase as a gift--bubble bath	1	21.0	11.7**	21.9	12.1	8.1**	17.1
<hr/>							
Purchase for Self vs Purchase as a gift--blanket	1	12.6	11.4**	6.1*	14.3	8.5**	4.9*
<hr/>							
Purchase for Self vs Purchase as a gift--record album	1	10.3**	5.0*	.3NS	5.7*	1.4NS	.1NS
<hr/>							
Purchase for Self vs Purchase as a gift--jeans	1	.ONS	.3NS	.6NS	.8NS	.ONS	.1NS

\*probability less than .05  
\*\*probability less than .01  
all others (except NS) probability less than .001

FIGURE 1

CELL MEANS OF DEPENDENT VARIABLES



The third hypothesis, that there would be no interaction effect between task involvement and product involvement, however, must be rejected. For all of the real choice measures and for the relative choice measure for time expenditure, the interaction effect is significant. As may be seen in the plots of cell means and in the simple main effects tests, the effect of task involvement was to increase the amount of anticipated effort for the low involvement products, but not for the high involvement products (with the partial exception of the amount of money and number of stores for the record album).

Discussion

The one surprising finding is the interaction effect between inherent product involvement and task involvement. Although the stipulation that a product is being selected as a gift generally increases the reported amount of consumer effort expenditure, this is not the case for the high involvement products. The most plausible explanation is that there is a ceiling effect on the overall amount of involvement. For selection of a product like jeans, involvement is already as high as it can normally get and the gift conditions can raise this level no higher. On the other hand, gift-giving is able

to transform an otherwise low involvement purchase like bubble bath into an important decision. However, the possible existence of a ceiling effect on involvement need not mean that there is a comparable ceiling effect on the levels of arousal corresponding to this involvement. It is reasonably clear that even giving high involvement products like jeans or records as a gift is still not sufficient to create arousal at the extremely high levels for which decreased purchase effort is predicted.

Another way to view the combined effects of the levels of product involvement and task involvement studied, is in terms of product relevance to the purchaser's self concept. Compared to the two low involvement products (bubble bath and blankets) the two high involvement products (record albums and jeans) are relatively visible to others and also more likely to be construed as matters of "taste". However, as has been noted by Belk (1976), normally non-visible products tend to become visible symbolic communications when they are given as gifts. A part of this symbolic communication involves a message about the recipient (i.e. this is a gift that suits you), and a part of the communication involves a message about the giver (i.e. this is a gift that reflects my tastes). Although the addition of the message about the recipient might seem to justify greater effort expenditure for the inherently high involvement products, the fact that the normal level of effort expenditure for these products is already high enough to consider the symbolic aspects of the purchase, may mitigate this effect.

There are some obvious cautions which need to be applied to interpreting and generalizing these results. Since these are reports of anticipated effort, they are subject to social desirability and other biases and may not reflect the actual amount of effort devoted to purchases. It is, however, encouraging to note that Ryans (1977) found greater amounts of information search time were expended when small appliances were chosen as gifts for people outside of the giver's household than when they were chosen for personal use. Nevertheless, the theoretical interpretations of these effects as due solely to task involvement are tentative given the singular manipulation of this variable. The same is true of the product involvement variable since the products employed also differ in other ways such as their visibility. Also, although the products selected were relevant and differentially involving for the college female sample, they may not be so for others.

#### Conclusion

It appears that the reason for purchasing an item as well as the inherent involvement with the item create differences in the amount of effort a consumer is willing to exert in purchasing the product. It would therefore be a mistake to assume that personally uninvolved products are purchased with little effort devoted to securing and processing information or otherwise attempting to optimize purchase selection. In fact, during peak holiday gift-giving periods high task involvement and attendant effort may predominate the selection of normally uninvolved products. As suggested by other research (e.g., Hupfer and Gardner, 1971) as well as the present research, it would also be a mistake to assume that higher price is necessarily related to higher involvement and higher purchase effort. Even though blankets are generally more expensive than jeans, jeans are higher involvement products to which greater amounts of purchase effort are devoted.

Among the areas for future research suggested by this study are investigations of purchase situation characteristics, other than personal/gift use, which affect

the level of task involvement and purchase effort. It is also of interest to know whether different bases for product involvement (e.g. social relevance, financial relevance, hedonic relevance) are all able to create greater purchase effort expenditure. It will also be necessary to establish these effects in actual as well as in anticipated purchase effort expenditures. And the particular types of purchase effort expenditure which are favored in different types of high involvement purchase decisions might be considered. Finally, it would be useful to know whether increases in consumer purchase effort in response to greater purchase involvement, have the same effects on postpurchase processes such as dissonance reduction as when the effort increase is necessitated by unforeseen purchase conditions.

#### References

- Russell W. Belk, "Situational Variables and Consumer Behavior," Journal of Consumer Research, 2(1975), 157-64
- Russell W. Belk, "It's the Thought that Counts: A Signed Digraph Analysis of Gift-Giving," Journal of Consumer Research, 3(1976), 155-62
- Russell W. Belk, "Gift-Giving Behavior," in Jagdish Sheth (ed.), Research in Marketing, Vol. 2, (Greenwich, Conn.: JAI Press, forthcoming)
- Daniel E. Berlyne, "Arousal and Reinforcement," in D. Levine (ed.), Nebraska Symposium on Motivation, (Lincoln, Neb.: University of Nebraska Press, 1967)
- Richard N. Cardozo, "An Experimental Study of Customer Effort, Expectation, and Satisfaction," Journal of Marketing Research, 2(1965), 244-49
- Richard N. Cardozo, and Dana Bramel, "The Effect of Effort and Expectation on Perceptual Contrast and Dissonance Reduction," Journal of Social Psychology, 79 (1969), 55-62
- Kjell Grønhaug, "Buying Situation and Buyer's Information Behavior," European Marketing Research Review, 7 (1972), 33-48
- Flemming Hansen, Consumer Choice Behavior: A Cognitive Theory (New York: Free Press, 1972)
- Donald O. Hebb, A Textbook of Psychology, 2nd ed. (Philadelphia: Saunders, 1966)
- John A. Howard and Jagdish Sheth, The Theory of Buyer Behavior (New York: Wiley, 1969)
- Nancy T. Hupfer and David M. Gardner, "Differential Involvement with Products and Issues: An Exploratory Study," in M. Venkatesan (ed.), Proceedings, 2nd annual conference of the Association for Consumer Research, (University of Maryland, 1971, 262-70)
- George Katona and Eva Mueller, "A Study of Purchase Decisions," in L. H. Clark (ed.), Consumer Behavior, (Vol. 1, New York: NYU Press, 1954)
- John L. Lastovika, "An Exploratory Multidimensional Scaling Study of Involvement with Products and Services," mimeographed paper (University of Illinois, 1976)
- Joseph W. Newman, "Consumer External Search: Amount and Determinants," in A. G. Woodside, J. N. Sheth, and P. D. Bennett, Consumer and Industrial Buying Behavior, (eds.), (Amsterdam: Elsevier North-Holland, 1977)

Brian T. Ratchford and Alan A. Andreasen, "The Effects of Role Structure, Decision Type, and Household Characteristics on Consumer Information Source Use Under Conditions of Limited Experience," Working Paper #181, (University of Illinois, College of Commerce and Business Administration, 1974)

Michael L. Rothchild, "Advertising Strategies for High and Low Involvement Situations" paper presented at 1977 Attitude Research Conference, Las Vegas, Nevada

Adrian B. Ryans, "Consumer Gift Buying Behavior: An Exploratory Analysis," in B. A. Greenberg and D. N. Bellenger, Contemporary Marketing Thought (Chicago: American Marketing Association, 1977, 99-115)

Musafer Sherif and H. Cantril, The Psychology of Ego Involvement (New York: Wiley, 1947)

M. Venkatesan and Johan Arndt, "Consumers' Use of Time: An Exploration with Time Budget," in Jagdish N. Sheth (ed.) Research in Marketing, (Vol. 3, Greenwich, Conn.: JAI Press (forthcoming))

Arch G. Woodside, "Positive Disconfirmation of Expectation and The Effect of Effort on Evaluation," Proceedings, 80th Annual Convention of the American Psychological Association (1972), 743-44

GIFT-GIVING:  
A REVIEW AND AN INTERACTIVE PARADIGM

Sharon K. Banks (student), University of Oregon

Abstract

This paper consolidates past theories and findings on gift-giving with the limited contemporary work done on gift-giving, primarily within the marketing discipline. Further, an interactive gift-giving model, including a view of both the giver and receiver, is presented, with which future research directions are suggested.

Introduction

It has been conservatively estimated that gifts account for 10% of all retail sales in North America (Belshaw, 1965). Further, many retailers estimate that 30 to 50% of their total volume (and even higher percentages of their total profits) occur in two months of the year, November and December. These astounding figures are due primarily to heavy gift-buying during the holidays. To further indicate the extent of gift-giving, in a 1973 study of 219 gift-giving instances by 73 Philadelphia area residents, Belk (1977) found an average of 26-27 gifts were purchased during the previous year per adult with a total average annual expenditure of \$280.43.

It has been suggested that marketers should aim at encouraging greater gift-giving and developing more gift-giving opportunities (Lowe, et. al., 1971). This is particularly relevant due to the reciprocal nature of gifts, i.e., almost every sale generated by primary promotional activity leads to two sales actually being made. Despite all the above compelling reasons to study gift-giving in depth, marketers have steadfastly ignored the subject almost entirely. Only a few researchers in marketing have shown any interest in the topic.

Most of the early work done on gift-giving arises from studies by anthropologists or sociologists who have examined gift-giving within primitive societies (Mauss, 1954, and Malinowski, 1932). The concepts involved in these early studies were then applied to social exchange theory among sociologists such as Levi-Strauss (1965), Homans (1950, 1965), and Blau (1965) and in the study of gratitude by psychologists such as Jones (1964).

In the sixties the study of gift-giving branched out widely. Schwartz (1967) studied the social psychology principles involved in gift-giving. Titmuss (1971) looked at the gift of giving blood. Others, such as Leeds (1963) and Krebs (1970) approached gift-giving from the standpoint of altruism.

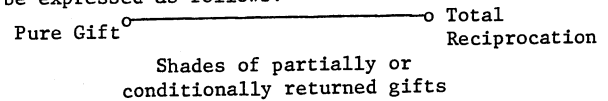
There is a need to consolidate past theories and findings on gift-giving with the contemporary work done on gift-giving. This paper attempts this task while also developing an interactive gift-giving model and suggesting future research directions for the field.

Gift-Giving's Relationship to Behavioral Concepts

A useful approach for reviewing the literature on gift-giving is to focus on the behavioral concepts which have been studied in conjunction with gift-giving. The major concepts that have been linked with gift-giving include reciprocity, interaction, and identity formation.

Gift-giving and Reciprocity

One way to view gift-giving is according to the reciprocity commitments involved with the gifts. Lowe, et. al. (1971) has set up such a classification scheme. It can be expressed as follows:



A pure gift would be an altruistic act where nothing is given or expected in return. Pure gifts were not often observed by researchers of primitive societies. Mauss (1954), and others found that gifts were conceived as socially meaningful acts.

Although not stressed by the early writers on gift-giving, altruism is a relevant motive. It would appear that this motive often arises when the receiver is capable of appreciating the gesture but is incapable of returning the gift or favor, in particular, the very old, the very young, the very sick and the very poor. Leeds (1963) defines this process as the "norm of giving" in contrast to Gouldner's (1960) "norm of reciprocity."

On the other end of the continuum is total reciprocity, which involves the social obligations to give, to accept, and especially to repay (or reciprocate). Moral enforcement of these obligations is a concept addressed by most of the early authors on gift-giving, such as Mauss (1954), Simmel (1950), Levi-Strauss (1965), Homans (1965), Schwartz (1967), and Titmuss (1971). Gouldner (1960) indicates that reciprocity is expected in the form of a return gift or in feelings of gratitude or deference. Because of the wide range of repayments possible, Gouldner's reciprocity can be translated into the reciprocity continuum earlier depicted with gratitude classified as less reciprocation than deference and deference as less reciprocation than a return gift.

However, reciprocation may be mediated by other variables, such as the ability of the individual to return a gift and the differential position of the individual in society. A good example of this is Muir and Weinstein's (1962) study on the giving of favors and the social debt incurred by different social classes by asking samples of low class and middle class members questions about favors. Middle class statements indicated that social obligation required them to make certain repayments and to form expectations for others to repay their obligations or face growing hostility, i.e., the norm of reciprocity. On the other hand, lower class individuals tended to feel more "grateful" than "obligated," i.e., one gives when one is able, expecting others to do the same.

Gift-Giving and Interaction

Mauss (1954) looked at gift-giving within primitive societies and its prevalence caused Homans to use Mauss' (1954) work as a basis for his social exchange theory. Mauss looked at exchange of gifts in primitive societies and called this process "a total social fact" including social, religious, economic, utilitarian sentimental, legal and moral significance. Blau (1965) also began his theoretical treatise on the exchange of social rewards based on gifts in the form of deeds, which are expenditures of time and effort, being ex-

changed.

Belk (1977a) characterizes gift-giving as a process of symbolic communication with the gift being both the message and the channel. There are opportunities for both encoding errors (choosing proper gift) and decoding errors (understanding the gift's meaning). An additional complication is the possibly unreliable feedback due to simulated approval.

It has been suggested that those to whom an individual gives gifts are in some way different from those to whom no gift is given (Schwartz, 1967). Thus, the gift exchange process tends to dramatize and clarify group and interactive boundaries.

Further, gifts help to define an individual's status or status change in society, e.g., graduation gifts and Father's Day presents. Gifts also act as symbols of social support in the commonly recognized rites of passage, such as engagements, baby showers, religious confirmations and weddings. (Belk 1977a)

#### Gift-Giving and Identity Formation

Identity formation, which is synonymous with the development of self-concept, can be viewed from two perspectives: (1) the socialization of children and (2) the generation and development of identity in adults.

Parents, other family members and friends influence the socialization of a child with gifts. These gifts affect the child's developmental processes, interests, knowledge acquisition, view of the world and basic values. Gift influences are particularly crucial for pre-school children because the child contributes little or no input into the gift decisions.

The process of socialization includes the learning of sex-role expectations. It is reasonable to presume that some sex-role stereotyping occurs simply because the "correct" sex-role gift is given. Belk (1977b) examined the possession and perception of various sex-role related toys by a small sample of 22 preschool children (4 and 5 year olds) and their parents. He arrived at a number of interesting conclusions, indicating the significance of gifts in the socialization of children: (1) As expected, children own more same sex toys than cross-sex toys. (2) However, the above pattern is weaker for toys given by the child's parents. (3) Toy ownership plays a significant role in the development of sex-role preference.

Caron and Ward (1975) conducted a study in Canada on the interaction of children and parents involving gift decisions and reached the following conclusions: (1) Almost  $\frac{1}{2}$  of the Christmas gifts purchased by parents are not explicitly requested by the children. (2) The types of gifts requested varied by social class. (3) Parents mediate gift requests, i.e., they selectively yield to purchase requests.

The socialization process is simply the first step in a long progression of the gift as a generator of identity. Gifts are one of the ways in which the pictures that others have of us in their minds are transmitted (Schwartz, 1967). Thus, the giver expresses his perception of the recipient's self-concept. But, further, gifts also communicate something about the giver. That is, individuals confirm who they are by what they give.

#### Gift-Giving As Viewed Within A Marketing Context

Although there has not been extensive published research on gift-giving within a marketing context, there have been several studies which have produced some interest-

ing results. These studies can be categorized under three main topic areas: (1) Psychological interrelationships; (2) Risk reduction and information sources used; (3) Gift context studies in which the dimensions of gift-giving are investigated.

#### Psychological Interrelationships in Gift-Giving

Belk (1976), utilizing Heider's balance theory principles, investigated to what extent giver traits and preferences versus giver perceptions of recipient traits and preferences determined the gift selection. The diagram is based on positive/negative relationships between the giver's self-image, giver's liking of recipient, giver's evaluation of gift and giver's perception of recipient's affect toward the gift. Different combinations, such as all the above positively related, will lead to a state of balance (or imbalance).

Belk measured the above concepts on a sample of gift-givers and ascertained that there were a significant number of balanced configurations but that balance does not explain the entire concept. Where imbalance occurred, satisfaction with the giver's choice was much lower. Furthermore, imbalanced situations were more likely to occur when the giver and recipient had not established a prior history of reciprocal giving and when the recipient was not a close member of the giver's family.

Through the use of canonical analysis, Belk (1977a) ascertained that the ideal self-concept of the giver was more highly correlated with the choice of the gift than was the giver's self-concept and perceptions of the recipient, although these latter two were also found to be significant factors.

#### Perceived Risk and Information Search in Gift-giving

Hart (1974) found that subjects were more conservative when making decisions for their spouses than when making the same decisions for themselves. Subjects rated the overall perceived risk severity as much greater when deciding for their spouses than for themselves.

In studying the purchase of a kitchen knife for home use versus as a gift, Vincent and Zikmund (1975) found that social risk was perceived as much greater when an item was purchased as a gift versus when it was purchased for home use.

However, the opposite results were obtained with financial risk, i.e., it being significantly higher for home use than as a gift purchase. Purchasing a more expensive model for a wedding gift than for personal use could be interpreted as a risk reduction method for handling perceived social risk. Financial risk would thus not be considered as important to the purchaser. Likewise, Shapiro (1970) found that "when buying a gift, respondents viewed the quality of the product as more important and the price as less important than when buying a product for herself."

In a simulated shopping trip, Weigl (1975) used the same risk measures as Vincent and Zikmund. His results on financial risk supported Vincent and Zikmund's findings. However, he did not find any significant differences with social risk when using the multiplicative relationship between uncertainty and consequences (components of risk). He then tested each component separately and found that gift purchases were significantly more concerned about the "consequences" of a bad experience with the product than non-gift purchasers, whereas the uncertainty dimension was not significantly different for the two groups.

Additionally, contrary to Weigl's expectations, there



were no significant differences found for either the extent of information search or the time spent in this search activity. However, Gronhaug (1972) found that buyers of tableware utilized different types and sources of information depending upon whether the purchase was for personal use or for a gift.

Gift purchasers have also been found to be more likely to begin shopping with a pre-specified target price range than those buying the same items for personal use (Ryan, 1977). Belk (1977a) suggests that this is because gift-giving participants are seeking fair exchange.

#### Gift Context Studies

Because gift-giving has been and is still being approached from so many diverse directions, there must be a clear delineation made of the various dimensions of gift-giving which should be of interest to marketers. These dimensions have been developed primarily by a review of Belk's enumeration of dimensions and research efforts and Lowes, Turner and Wills' review (1971) of four important marketing studies on gift-giving, including three British studies (Bradford study, Gallup study, National Opinion Poll study) and an American study, all conducted in the mid-sixties.

Occasion. The frequency of all gift-giving occasions as reported by Belk's 1973 Philadelphia sample is indicated in Table 1. Bussey (1967) reports similar findings for the two top categories in a British study except in reverse order of prevalence. Then, cultural differences begin to show up, with the third most prevalent gift-giving occasion being holiday trips in England.

Belk (1977a) hypothesized that the price of gifts bought for close family members would be higher than for all others. This was found to be true but the relationship was not strong. However, when examined by occasion, the relationships were much stronger, with Christmas, weddings and anniversaries providing occasions for more expensive gift-giving (Table 1).

TABLE 1  
FREQUENCY OF GIFT-GIVING BY OCCASION AND PRICE RANGES

Occasion	Frequency Percent	Price of Gift-Percent	
		Less Than \$10	More Than \$10
Birthday	35	51	49
Christmas	29	39	61
Wedding	7	7	93
Mother's Day/ Father's Day	6	54	46
Wedding Anniversary	5	27	73
Graduation	5	40	60
No Occasion/Other	12		

Relationship Between Giver and Receiver. Data from the Bradford study was used to compile a reciprocity coefficient to indicate the degree of reciprocation between various parties in the gift-giving situation (Lowes, et. al., 1971). This ratio was simply:

$$\frac{\text{No. of presents received from}}{\text{No. of presents given to}}$$

It was below 1 for children, in-laws, father, unmarried brothers, near 1 for spouses, mother, uncle, neighbors, and greater than 1 for business contacts, friends, and workmates.

Further, Belk compiled a table of gift recipients by the frequency of gifts received. This can be compared to a similar table which can be compiled from the information provided from the Bradford study, after Christmas gifts had been eliminated. The closeness of the percentages, when comparable, are phenomenal considering the variations in the two studies.

TABLE 2  
FREQUENCY OF GIFT RECIPIENTS REPORTED

Recipient	Percent Philadelphia Study	Percent Bradford Study
Friend	33	?
Parent/Parent-in-law	20	19
Child/Child-in-Law, Niece/Nephew	18	18
Sibling/Sibling-in-Law	12	13
Spouse	8	8
Other	9	42 (no classification given for friend)

Type of Gift Selected. Another important element in looking at gift-giving is the type of gift selected. Belk's Table of Frequency of Gifts Reported is reproduced here as Table 3. It is interesting to note that the clear favorite is clothing. This same result was reported in the English Gallup Poll survey on Christmas gifts.

The Bradford study tried to categorize the most relevant types of gifts given by occasion. It indicated that personal gifts were the most popular gifts at Christmas, followed by novelties and then household items. Weddings and engagements were characterized by household gifts. Personal gifts were predominant on birthdays, anniversaries and Mother's/Father's Days. Money gifts came into significant use only on birthdays.

Other Descriptive Information. Belk described some interesting descriptive results of his gift-giving study: (1) Gift-giving appears to be a pleasurable act for most people over most occasions, (2) Gift selection was often aided by others (38% of sample received hints, requests, or were aware of recipient's desires, (3) The shopping experience for gift purchases was characterized as very easy or fairly easy by almost 2/3 of the sample; less than 1/3 required more than 1 hour shopping time.

One final interesting finding noted in the Bradford study was an attempt to enumerate the primary reasons for giving (Table 4). There was no attempt in this study to relate the reasons below to specific occasions, which would have provided some additional insights. As Lowes, Turner and Wills suggest, most of the below occasions are happy occasions. Once these

occasions have become established, a social norm for gift-giving compels one to conform, thus, the rather high percentage of people answering "expected thing." There is still much left to be examined within gift-giving behavior. However, these preliminary studies give an excellent indication of the direction that future descriptive studies should follow.

TABLE 3  
FREQUENCY OF GIFTS REPORTED

Gift	Number of Reports	Percent
Clothing	54	24.66
Jewelry	23	10.50
Sporting goods	13	5.94
Personally made item	12	5.48
Phonograph Record/Tape	10	4.57
Home furnishing accessory	9	4.11
Appliance (minor)	7	3.20
Cologne/Perfume	6	2.74
Dinnerware/Houseware	6	2.74
Electronic Entertainment Equipment	6	2.74
Book	5	2.28
Cash/Stock/Bond/Gift Certificate	7	2.28
Craft Kits/Equipment	5	2.28
Flowers	5	2.28
Home Maintenance equipment	5	2.28
Keepsake (e.g., music box)	5	2.28
Plant	5	2.28
Alcohol	4	1.83
Clothing Accessory	4	1.83
Game or toy	4	1.83
Linen	4	1.83
Novelty (e.g., poster)	4	1.83
Luggage	3	1.37
Personal care product	3	1.37
Other	12	5.48
<b>Total</b>	<b>219</b>	<b>100.0%</b>

TABLE 4  
REASONS FOR GIVING

To obtain pleasure	27%
Show friendship/love	16
Expected thing	16
To give pleasure	13
To show appreciation	13
Sentimentality	5
Other	6
Don't know	4

Gift-giving - Future Research Directions

There are two different ways to approach the study of gift-giving, viewing gift-giving as either a situational variable or as the primary variable of interest. Descriptive studies, such as Belk's 1973 study and the

three British studies, plus Belk's work on balance theory, self-concept and socialization and Ward and Caron's socialization paper are of the latter nature.

Gift-Giving as a Situational Variable

However, some of the other studies, such as Weigl's perceived risk, Gronhaug's information sources and Ryan's appliance study, really used gift-giving as a situational variable and looked primarily at how gift buying differed from non-gift buying.

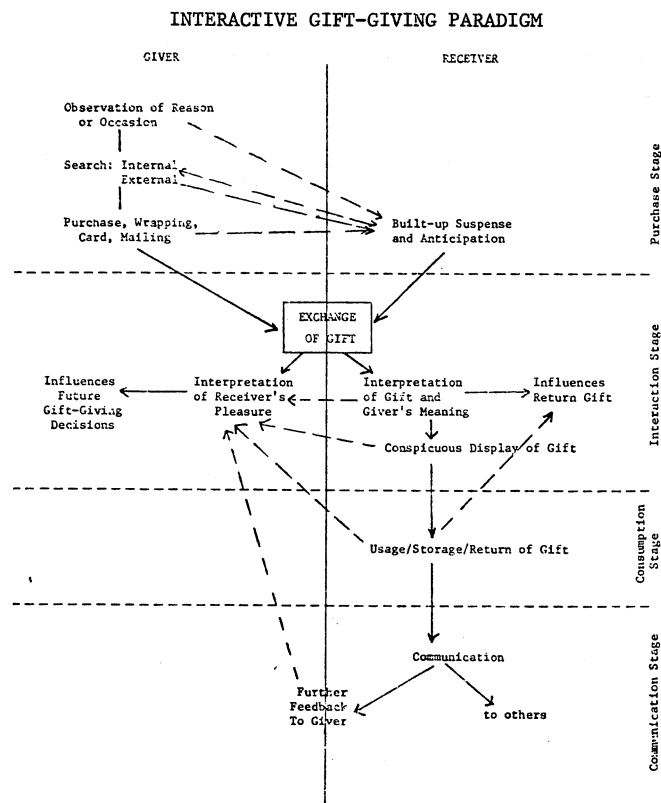
Gift-giving is a rich area to be incorporated into any study of consumer behavior where situational variables will increase the accuracy of predictions. Gift-giving has been found to be related to risk perception, information sources and expected price ranges used. Further, several researchers have suggested (based on no empirical findings) that a gift-buying situation increases the importance of the purchase (Granbois, 1963 and Katona and Mueller, 1955). Thus, gift-giving situations probably increase the level of involvement in the purchase decision, which has a number of important implications for marketing.

An Interactive Gift-Giving Paradigm

Gift-giving pervades through all levels of consumer behavior. A beginning approach to gift-giving requires a four-stage process including: (1) Purchase Stage, (2) Interaction/Exchange Stage, (3) Consumption Stage, (4) Communication/Feedback Stage.

Figure 1 models the four stages, their component parts and the various interrelationships between the two primary parties, the giver and the receiver. Because gift exchange is often two-sided, the giver may also be a receiver and vice versa during the same time interval. Or the exchange may be one-sided and due to the interpretation of the first gift coupled with the observation of a need or occasion the process is often replicated (reciprocity, following some time lag) with the giver becoming the receiver and vice versa.

FIGURE 1



The interactive paradigm suggests that there are two parties in the process, rather than just one as normally visualized in consumer behavior models. The purchase stage includes the giver's observation of a reason or occasion, internal and external search and the actual purchase including subsidiary activities (gift wrapping, etc.). At all of these points, the receiver may be building up suspense and anticipation and may have even aided the giver by giving him hints or gift suggestions.

During the interaction stage, the gift is exchanged. It is at this stage that the receiver interprets the gift and its meaning and the giver interprets the receiver's pleasure in the gift. Both of these interpretations strongly influence the return gift and future gift giving. Further, the receiver may conspicuously display (wear, use) the new gift immediately, which allows for further interpretation by the giver.

The consumption stage can be very simply viewed as the usage or storage of the gift, which further influences the receiver's return gift and the giver's interpretation of the receiver's pleasure. Finally, the communication stage involves the receiver communicating to others and to the giver his satisfaction or dissatisfaction with the gift.

Purchase Stage. This stage includes the giver becoming aware of either an underlying reason to give or an upcoming occasion, that is generally associated with gift-giving, or a combination of the two. Both specific reasons and specific occasions have been catalogued but not cross referenced in earlier studies. Initial studies have investigated what is purchased and price ranges by occasions.

The search process has only been touched upon by past studies. These studies indicate that a large percentage of gift purchasers get hints or suggestions, use their ideal concept primarily in terms of gift selection, perceive greater risk consequences, use different sources of information, but the search process is relatively easy and does not seem to involve more information search than non-gift purchases.

If one presumes that the majority of Udell's (1966) sample were gift-purchasers (implied but never specifically defined as such), then from this study of small appliance purchases shortly before Christmas, the following summarizations can be made about gift buyers' behavior patterns (in regards to one type of gift purchase): (1) Most of the purchases (73%) were planned prior to shopping in a store; 13% after visiting a retail store. (2) Nearly 60% of the respondents had shopped for the small appliance only in the store where the purchase was made.

The above would suggest gift purchases are highly planned as far as type of item to be purchased. However, this does not suggest that purchasers had planned their brand and model. But we still do not have information on what cues are used, which product attributes are the most relevant and how important brand image is.

The above conclusions also suggest that gift purchasers rely primarily on one store. Store image, as well as brand image, are both conspicuous aspects that give the gift receiver cues as to the value the giver is trying to impart. If store image is a very important cue (due to gift box, return policies or just the buyer's awareness that the recipient may need to return the item) then shopping behavior may be drastically affected. For example, many previous studies have indicated that individuals tend to shop in stores of equivalent social class (Martineau, 1958). Does this

change when the purchaser is buying a gift for someone in a different social class or where the purchaser wishes to impress the receiver, even in the same social class, by a prestigious store name?

Price, another important attribute in gift purchase behavior, seems to be relied on heavily as an indicator of quality, and expected price ranges are not unusual according to earlier studies. However, this is a complex topic which needs much more exploration.

The next step after search is the purchase and the gift wrapping procedure, card purchase, mailing procedure and anything else that leads up to the point of exchange. These procedures really involve the store's services. How important is gift wrapping, mailing, delivery and lenient return policies on the gift purchaser's decision to shop one store rather than another?

While the giver is progressing through all these various steps, anticipation or suspense is often involved with the receiver. Different levels of anticipation will be aroused dependent on the receiver's awareness of the giver's actions and previous gift-giving experiences with the giver. Gifts are kept secret or hidden for the sake of the giver as well as for the receiver because of the cruciality of the recipient's response to the gift (Schwartz, 1967).

Interactive/Exchange Stage. Schwartz suggested that the gift-giving occasion may be insulated by other activities, such as a Christmas dinner, so as to deemphasize the sudden release of built-up suspense. Belk (1977a) discussed the one-sided versus two-sided exchange procedure and the communication aspects of the gift exchange. Thus, the receiver interprets the meaning of the gift and this may affect his subsequent gift-giving behavior with the giver and his attitudes toward the giver.

The giver, realizing that social obligation may cause the receiver to feign pleasure in the gift, looks for whatever relevant cues are available to determine how the receiver felt about the gift. Thus, the receiver's reaction to the gift will be carefully observed by the giver. The nature of the gift will dictate what additional cues will be used. For example, can the product be immediately consumed (such as candy) or must it be conspicuously displayed or worn (such as a picture, vase or clothing)?

Consumption Stage. This stage includes the usage and/or storage of the gift. Was the gift returned because it did not adequately meet the receiver's needs? What types of gifts are returned most often and what are the predominant reasons for these returns? How does dissatisfaction in product use relate to gifts? Satisfaction has been found to be affected by expectation levels (Cardoza, 1965). The question then becomes how is expectation level related to gift anticipation? Are levels higher due to the surprise nature of gifts or lower because the item was not personally selected?

Communication/Feedback Stage. The final stage in the gift-giving process occurs after the receiver has used the gift for a period. Often, the receiver gives additional feedback to the giver. The receiver may express his satisfaction with the product after usage. The giver may, consciously or unconsciously, look for further reinforcement that the gift was satisfactory. The classic example of this is the display of disliked gifts put up just before arrival of out-of-town parents or relatives, who are sure to look for the display of these articles. The items are then removed again as soon as the visitors drive off. This type of untruthful but tactful behavior clouds perceptions held by parties towards the gifts. This continuing feedback

will affect the interpretation the giver makes of the receiver's satisfaction of the product and will, in turn, affect the giver's self-perception and attitudes towards the product, brand, store purchased in, the information source used, etc.

#### Summary

This paper has been an attempt to consolidate the literature on gift-giving and to suggest a simple paradigm for the continuing study of gift-giving within consumer behavior involving many implications for marketing management. Hopefully, this consolidation and approach to gift-giving behavior will be able to stimulate readers to pursue the many avenues of research still left totally unexplored.

Gift-giving is a pervasive form of consumer behavior engaged in on a frequent basis by all members of modern society. Gift-purchasing, gift-exchange, gift-consuming and gift-communication are separate but overlapping areas in which future research may explore gift-giving behavior. Only by consciously studying and conscientiously describing gift-giving behavior will important relationships between variables be ascertained and applied to marketing management.

#### References

- R.W. Belk, "It's the Thought that Counts: A Signed Digraph Analysis of Gift-Giving," Journal of Consumer Research, 3 (December 1976), 155-62.
- \_\_\_\_\_, "Gift-Giving Behavior, Part A," Faculty Working Papers, College of Commerce and Business Administration, University of Illinois at Urbana-Champaign, 1977a.
- \_\_\_\_\_, "Gift-Giving Behavior. Part B," Faculty Working Papers, College of Commerce and Business Administration, University of Illinois at Urbana-Champaign, 1977b.
- C.S. Belshaw, Traditional Exchange in Modern Markets (Englewood Cliffs, N.J.: Prentice-Hall, 1965)
- P. Blau, "The Exchange of Social Rewards," in L.A. Coser and B. Rosenberg, eds., Sociological Theory, (New York: Macmillan, 1965)
- J. Bussey et al., "Patterns of Gift-Giving," Management Centre, Bradford University, 1967.
- A. Caron and S. Ward, "Gift Decisions by Kids and Parents," Journal of Advertising Research, 15 (1975), 15-20.
- A. Gouldner, "The Norm of Reciprocity: A Preliminary Statement," American Sociological Review, 25 (1960), 176-7.
- Donald H. Granbois, "The Role of Communication in the Family Decision-Making Process" in S. Greyser, ed., Toward Scientific Marketing (Chicago: American Marketing Association, 1963)
- K. Gronhaug, "Buying Situation and Buyer's Information Behavior," European Marketing Research Review, 7 (1972) 33-48.
- E.W. Hart, "Consumer Risk-Taking for Self and Spouse," unpub. Ph.D. dissertation, Purdue University, 1974.
- G. Homans, The Human Group (New York: Harcourt, 1950)
- \_\_\_\_\_, "Social Behavior as Exchange," in L.A. Coser and B. Rosenberg, eds., Sociological Theory (New York: Macmillan, 1965)
- E.E. Jones, Ingratiation (New York: Appleton-Century-Crofts, 1964)
- G. Katona and E. Mueller, "A Study of Purchasing Decisions," in L.H. Clark, ed., Consumer Behavior: The Dynamics of Consumer Reaction, (New York: University Press, 1955, 30-87)
- D. Krebs, "Altruism: An Examination of the Concept and a Review of the Literature," Psychological Bulletin, 73 (1970), 258-302.
- R. Leeds, "Altruism and the Norm of Giving," Merrill-Palmer Quarterly, 9 (1963) 224-39.
- C. Levi-Strauss, "The Principle of Reciprocity," in L.A. Coser and B. Rosenberg, eds., Sociological Theory, (New York: Macmillan, 1965)
- B. Lowes, J. Turner, and G. Wills, "Patterns of Gift-Giving," in G. Wills, ed., Exploration in Marketing Thought (London: Bradford University Press, 1971, 82-102)
- B. Malinowski, The Sexual Life of Savages in North-Western Melanesia (London: Routledge, 1932)
- \_\_\_\_\_, "The Principle of Give and Take," in L.A. Coser and B. Rosenberg, eds., Sociological Theory (New York: Macmillan, 1965)
- Pierre Martineau, "Social Classes and Spending Behavior," Journal of Marketing, 23 (October 1958), 121-30.
- M. Mauss, The Gift: Forms and Functions of Exchange in Archaic Societies, English translation by F. Cunnison (Glencoe, Ill.: Free Press, 1954)
- D. Muir and E. Weinstein, "The Social Debt: An Investigation of Lower Class and Middle Class Norms of Social Obligation," American Sociological Review (1962), 532-9.
- A. Ryans, "Consumer Gift Buying Behavior: An Exploratory Analysis," in D. Bellinger and B. Greenberg, eds., Contemporary Marketing Thought, Chicago: American Marketing Association (1977), Series No. 44, 99-104.
- B. Schwartz, "The Social Psychology of the Gift," American Journal of Sociology, 73 (1967), 1-11.
- B.P. Shapiro, "The Effect of Price on Purchase Behavior," in D.L. Sparks, ed., Broadening the Concept of Marketing (Chicago: American Marketing Association, 1970)
- G. Simmel, The Sociology of Georg Simmel, English translation by K. Wolff (Glencoe, New York: Free Press, 1950)
- R.M. Titmuss, The Gift Relationship (London: Allen and Unwin, 1971)
- J.G. Udell, "Prepurchase Behavior of Buyers of Small Electrical Appliance," Journal of Marketing, 30 (October 1966), 50-51.
- M. Vincent and W. Zikmund, "An Experimental Investigation of Situational Effects on Risk Perception," in M.J. Schlinger, ed., Advances in Consumer Research, (Chicago: Association for Consumer Research, 1975)
- K.C. Weigl, "Perceived Risk and Information Search in a Gift-Buying Situation," Masters Thesis, Purdue University, 1975.

## GIFT VERSUS PERSONAL USE BRAND SELECTION

Roger Heeler, York University  
June Francis, Proctor & Gamble Co. of Canada Ltd.  
Chike Okechuku, Bell Canada  
Stanley Reid (student), York University

### ABSTRACT

An information board study compared product data acquisition and choice in gift versus personal use scenarios. No difference in shopping behaviour was found when the gift was for a close friend. When the gift was for a "friend's wedding" less effort was devoted to the gift buying than to the personal use buying.

### INTRODUCTION

This study is concerned with behavioural differences between purchasing an item for personal use and purchasing it as a gift. Gifts form an important part of the total sales of some consumer products. For a marketer of a giftable product, it is important to know (a) how and why his product class may be chosen as a gift, and (b) how and why his brand may be chosen. Since marketing effort with respect to (a) will also benefit competition, (b) is likely to be more critical to the marketer and is the focus of this study.

The phenomenon of gift giving is in itself a subject interesting to psychology and sociology because of the complex motivations and symbolism involved. For example Schwartz (1967) noted that the gift object was more than merely the physical object -- it is a symbol of the giver's perception of the recipient, and the recipient's acceptance of the gift object amounted to his acceptance of the giver's idea of him. Levi-Strauss (1965) also saw gift-giving as serving needs beyond the mere exchange of commodities. He wrote:

"...goods are not only economic commodities but vehicles and instruments for realities of another order: influence, power, sympathy, status, emotion, and the skillful game of exchange consists of a complex totality of maneuvers, conscious and unconscious, in order to gain security and to fortify one's self against risks incurred through alliance and rivalry".

Thus in gift-giving it is not just "the thought that counts", but involved is a system of reciprocal maneuvering through which the giver and the recipient achieve higher order objectives.

Although the motivational and cognitive theories of gift-giving seem fairly well developed, not much attention has been paid to the behavioral aspects of the act. Since it has been shown that the gift object has evaluative dimensions vis-à-vis the recipient, and that the reciprocity relationship is characterized by a goal structure, it would seem that buyer behavior with respect to objects purchased as gifts would be different to that for objects purchased for own or normal household use.

#### Consumer Behaviour Research

Recently Ryans (1977) investigated the gift versus personal use purchases of small electrical appliances, using a recall questionnaire method. He found that appliances for gifting were more likely to be purchased from stores with quality image than those for

personal use. He also found that the amount to be spent was more likely to be set before the purchase, and that the time between the decision to purchase and the actual purchase tended to be shorter for gifts to members of a different household (outcome gifts). Generally he found no significant behavioural differences between purchasing for personal use and purchasing for a gift to a member of the same household (inhome gifts) with regard to price-setting and pre-purchase deliberation. Perhaps this was due to the greater cognitive balance between members of the same household (Belk, 1976).

Weigl (1974) investigated a hypothesis that gift giving would involve more perceived risk than purchases made for personal use. Weigl's product was tennis racquets and his gift scenario asked respondents to consider buying the product as a gift for a close friend of the same sex. His main hypothesis was not supported though a sub-hypothesis that perceived that financial risk would be greater for own use was supported. Weigl also hypothesized as a corollary that information acquisition and search time would be greater for gifts than for own use products. This hypothesis was also not proven. Weigl's report contains only the statistical significance or non-significance of the results so it is not known whether his results were in the hypothesized directions but not statistically significant at traditional levels of significance, or whether his results were contrary in direction to his hypotheses.

Heeler and Okechuku (1978) investigated data acquisition and search time for blenders in "own use" versus "gift for a friend's wedding" scenarios. They found significantly greater information acquisition for the "own use" condition. "Own use" respondents also took 30% longer to make decisions but this difference was not statistically significant at traditional levels of significance.

The studies described above appear to be in conflict with respect to whether gift buying is more, less, or equal in purchase activity to personal use shopping. But a resolution is possible if the contexts of gift giving are compared. Weigl found no difference when the gift was for a "close friend". Ryans found no difference for inhome gifts. Thus close friends and relatives appear to be treated equally with one's self in shopping. Heeler and Okechuku found less purchasing effort for a "friend's wedding". Ryans found less time was devoted to outhome gifts. Thus gifts to more distant people appear to involve less effort for the purchaser than own-use purchases. Perhaps this is why Belk (1976) in a study of gift satisfaction found that satisfaction with gift selection is likely to be higher under conditions of balanced cognitive configurations between the giver and receiver. Close friends or inhome recipients who are likely to be closer in cognitive configuration, have more shopping time spent on them which should result in greater gift satisfaction.

#### Hypothesis

It is hypothesized that in gift giving involving close

friends or relatives, as much trouble will be taken in the purchase as for an own use purchase. The combination of emotional bonding and the pressures to be perceived well in a close continuing relationship will lead to the recipient being noted as worth as much effort as one's self. By contrast a gift of a more distant, perhaps conventionally required, nature will involve less effort than for one's self. The purchasing duty must be adequately performed but is less involving than for one's self or a close friend or relative.

The information board method was extended to test this hypothesis as described below.

#### METHOD

To operationalize "distant" and "close" gift giving situations a "friend's wedding" and "a personal gift for a close friend" were chosen as alternate gift scenarios. Subjects consisted of 108 male and female students at a Polytechnic Institute. With ages between 20 and 30, and many enrolled as part-time students, the group was very familiar with weddings and close friends. Subjects received no monetary or participation points for participating in this study, but the Polytechnic teachers provided class time to the experimenters for a debriefing that used the exercise to illustrate the consumer behaviour research process.

Preliminary research on products suitable for the gift occasions led to two different products being used, blenders and watches. Blenders were appropriate to and used for "friend's wedding" but were inappropriate for a gift to a close friend. Watches were appropriate to and used for "close friend" but were inappropriate as a wedding gift. Thus product type and scenario are confounded in the research, but if the product and scenario were to realistically relate to each other this seemed to be inevitable. The real choice is between a personalized gift for a close gift giving situation, and a more utilitarian product for a more utilitarian occasion. A further constraint on product selection was product suitability for the information board procedure. For the procedure to have meaning there must be a good range of alternative brands in the product class, and a good range of attributes to differentiate brands. Watches and blenders both met this constraint. Blenders and watches were also products known to the respondents.

The Information Display Board (IDB) familiar to consumer behaviourists (for example see Berning and Jacoby, 1974) was used to elicit the depth, sequence, and content of information acquired prior to the choice of a brand of blender.

The (IDB) is essentially a board on which the attributes of the product class under study are displayed down the left hand column and the brands across the top (or vice versa) so that information on the brands is provided in a matrix format. The number of brands and attributes included determine the number of cells or bits of information in the matrix. Each cell contains the relevant brand-attribute information.

The information in the matrix cells is initially concealed from the subject. The subject is asked to make a product selection using as much or as little information from the board as they personally need. Thus the amount and type of information used by the subject is under his control, as it would be in a real shopping trip. The results of (IDB) studies usually show the same partial use of data in product selection as appear in observation and protocol studies of actual shopping trips.

Like other laboratory techniques, (IDBs) are artificial in requiring a special shopping simulation behaviour of subjects. Jacoby et al (1977) have found reasonably high correlations between (IDB) and verbal report data using breakfast cereals, margarine, and headache remedies, with respect to choice behaviour, thus providing some evidence for convergent validity (Heeler and Ray, 1972).

In this study, (IDBs) were used in an experimental design which should further improve the validity of the results. It was not own use or gift use purchasing which were interesting per se, but the difference between these two behaviours. Thus the artifact of laboratory experimentation was common to both experimental conditions and should net out when the difference is taken.

The boards for both products used sixteen brands and ten attributes for each brand. The sixteen watch brands were subdivided into eight male and female watches. Thus each board contained 160 data cells. Watch respondents operated in either the male watch or the female watch section and so effectively operated with 80 data cells. The data for each cell was mounted on the board, and concealed prior to respondent access by a card slipped into a simple holding rack. Brands were on the horizontal axis, attributes on the vertical axis. Brands and attributes were randomly assigned in order. The brand and attribute positioning was the same for all respondents. This results in a position effect (as described later), but the position bias was equal between the experimental conditions so no attempt was made to neutralize it.

#### Data Collection Procedure

The subjects were randomly assigned into four groups, a blender own use group (25), a watch own use group (27), a blender "gift for a friend's wedding" group (27), and a watch for "personal gift for a close friend" group (29). Each subject answered a product familiarity question and was then asked to choose a product according to the specific experimental group he was in. The subject was free to access as much or as little information as he deemed necessary to choose a brand and to announce his choice as soon as he made his decision.

When the subject indicated that he understood the instructions he was told to begin, and the experimenter activated a concealed stop watch. At the same time the interviewer was recording the codes identifying the bits of information accessed, in the sequence in which they were accessed. Data were collected on product familiarity, information bits selected and order of selection, time spent deliberating, and product chosen.

For watches there was a near equal proportion of women in both own use and gift conditions and high product familiarity for all respondents. For blenders there was also a near equal proportion of women in both conditions, but a higher proportion of gift subjects (82%) than own use subjects (56%) were highly familiar with blenders.

#### RESULTS

The principal hypothesis was supported. For the "close friend" scenario there was no significant difference in information accessed or time spent between the "gift" and "own use" conditions. For the "wedding" scenario there was significantly ( $t$  test  $p < .05$ ) more information accessed in the "own use" condition than in the "gift" condition. "Own use" respondents also took longer than the "gift" use respondents in making their selection, but this difference was significant only at the 15%

level. The detailed results are given in Table 1.

TABLE 1

	Blenders		Watches	
	Own Use	Wedding Gift	Own Use	Close Friend Gift
Number of information cells accessed	39.2	27.0	21.8	20.4
Search time (minutes)	6.0	5.0	4.6	4.2
Number of brands consulted	10.5	9.0	7.4	6.4
Number of attributes consulted	6.8	7.4	6.8	7.6
Transition type:				
II (same brand, different attribute)	.39	.31	.29	.43
III (same attribute, different brand)	.37	.31	.49	.27
IV (different brand and attribute)	.24	.38	.21	.30
Average price of product chosen	<u>\$33.1</u>	<u>\$34.9</u>	<u>\$29.9</u>	<u>\$31.8</u>
Sample size	25	27	27	29

In the methodology section it is noted that for blenders, more respondents were familiar with the product in the "gift" condition than in the "own use" condition. A rival hypothesis for the difference in data acquisition between the "gift" and "own use" condition could be that respondents familiar with the product would access less information, and since the "gift" respondents were more product familiar, this could account for the lower rate of accessing in the gift condition. In fact, in both "gift" and "own use" conditions, respondents more familiar with blenders accessed more information, so if no other factor were present one would expect a greater rate of information accessing in the gift condition, that is the opposite of the result obtained. In the "gift" condition "familiar" respondents averaged 28.3 data cells versus 21.4 for the "unfamiliar". In the "own use" condition "familiar" respondents averaged 44.0 data cells versus 33.0 for the "unfamiliar".

Table 1 also shows comparative results for brands and attributes accessed, and search transition types. It had been expected that the results for the number of brands and attributes accessed would parallel the main hypothesis, i.e. more for "own use" in the "wedding" scenario, equivalence for the "close friend" scenario, but this expectation was not borne out. For the "wedding" scenario, "own use" subjects accessed more brands than "gift" respondents, but this was only significant at the 12% level. "Wedding" scenario "gift" subjects actually accessed slightly more attributes than "own use" subjects. In the "close friend" scenario, "own use" subjects accessed more attributes but less brands than "gift" subjects.

In order to investigate possible differences in information acquisition sequence between purchasing situations use was made of the 4 different types of information search strategies identified by Jacoby et al (1975). Transition type 1 refers to a same brand-same attribute acquisition in a sequence of two information acquisitions, that is, immediate re-access of the same-brand attribute information. This transition was not possible in the present study due to the use of a fixed-board design. Type 2 transition refers to a same brand-different attribute sequence; type 3 to a different brand-same attribute sequence; and type 4 to a different brand-different attribute sequence.

The proportions of each respondent's information search involving types 2, 3 and 4 search strategies were calculated and reported in Table 1. For both "wedding" and "close friend" scenarios there was a higher proportion of type 4 transitions in the "gift" condition than in the "own use" condition, but this was not statistically significant.

Brands selected were similar for "gift" and "own use" conditions for both "wedding" and "close friend". This is illustrated by the average price of brand chosen data in Table 1. The blenders in the "wedding" scenario ranged in price from \$22 to \$43 with an average of \$34. The average price of blenders chosen was \$33 "own use" and \$35 "gift". For watches in the "close friend" scenario, prices ranged from \$12.95 to \$48 with an average price of \$34. The average price of blenders chosen was \$30 in the "own use" condition and \$32 in the "gift" condition. Thus slightly higher priced brands were selected in both scenarios for "gifts", but the differences were not statistically significant.

Table 2 shows the distribution of information access by attributes.

In the "wedding" scenario with blenders the most accessed attributes were price, brand name, and number of speeds. After allowing for the overall difference in information accessed, the differences in individual attributes accessed was small. Price was the most accessed attribute for "own use" but brand name was most accessed by "gift" subjects. "Gift" subjects also emphasized number of speeds and wattage more than "own use" subjects. These overt indicators, brand, speeds, and wattage, may be symbolically important in the gift situation.

In the "close friend" scenario with watches the most attributes accessed were consistent in both "gift" and "own use" conditions. Watch design (a picture of the watch), brand name, and price accounted for over half of the information accessed.

#### Methodological Considerations

While the (IDB) gave excellent control for a laboratory setting, and many of the likely biases were balanced out by the test and control group design, the authors were left with some concerns about the general applicability of the method.

The scenario itself requires subjects to take on a forced, "left brain" mock shopping task. In debriefing, most subjects admitted to some task uncertainty in the first rounds of data acquisition, and about 15% to experiencing confusion or difficulty with the task. Observation of subjects in process also suggested some of the "yeah saying" and "show I am logical" behaviour that is a problem in questionnaire research. Perhaps the most extreme case was a subject who announced at the start that "you will see what a fast decision maker I am" and then took twenty minutes longer than any other subject, carefully taking every piece of data from the board.

In debriefing subjects also reported some feeling of time constraint, despite the study's use of the standard (IDB) phrasing of "take as much or as little information as you like". This in fact may parallel true-to-life shopping. One respondent likened the feeling to going shopping with a friend and not wanting to keep the friend waiting too long. This feeling of time pressure was despite the subjects accessing more board information than in many other reported studies. The respondents accessed an average of 28 pieces, compared to 7 in Jacoby et al (1975).

TABLE 2  
INFORMATION ACQUISITION BY ATTRIBUTE

Attribute	Average		Average	
	Access Frequency	Percentage of Access	Access Frequency	Percentage of Access
<b>Wedding (blender)</b>				
Number of speeds	4.64	11.80	3.74	13.85
Safety of operation	3.52	8.99	2.85	10.56
Quietness of operation	4.16	10.62	1.92	7.13
Wattage	1.12	2.86	1.22	4.52
Brand name	5.80	14.80	4.77	17.70
Container material	3.08	7.87	1.85	6.86
Ice crushing option available	2.08	5.31	1.56	5.76
Warranty	4.48	11.44	2.30	8.50
Price	7.72	19.71	4.59	17.00
Ease of cleaning	2.56	6.53	2.19	8.09
<b>Close Friend (watch)</b>				
Water resistant	1.78	8.16	1.82	8.97
Manner of winding	1.07	4.93	1.59	7.78
Watch design	5.59	25.68	4.34	21.32
Warranty	1.00	4.59	1.86	9.14
Brand name	3.85	17.68	3.41	16.75
Day/date	0.74	3.40	0.93	4.57
Anti-magnetic	1.11	5.10	0.76	3.72
Price	3.07	14.12	2.82	13.87
Type of strap	2.03	9.35	1.48	7.28
Shock proof	1.56	7.14	1.34	6.60

In pilot work for the study a re-access (IDB) was tried. With the re-access board, subjects see a particular piece of information when they first access it, but to view it again they have to re-access it. This process caused great annoyance to respondents and breakdowns in realistic participation.

A strong Northwest corner effect was noticed in board use, that is respondents are biased to accessing information from the top left hand corner of the board. For example, the top eight brands of blenders were accessed 62% more than the bottom eight, and (combining male and female versions) the top four watches were accessed 55% more than the bottom eight. This effect is analogous to the inevitable shelf position effect on brands in a store but in a study that was not researching shelf effect could lead to misleading interpretations. There was a similar though less pronounced left side bias with attributes. In the current study these order biases were equal for the experimental conditions compared. In other research (in progress) the authors rotate the board position of brands and attributes between respondents.

#### DISCUSSION

The hypothesis given earlier stated that gift purchases for close friends or relatives will involve as much effort as personal shopping, but more distant gifts will involve less effort. The hypothesis was supported by the study in terms of information acquired. The lower information acquisition in the "wedding gift" condition was in terms of amount of information per brand or attribute, rather than in terms of number of brands or attributes considered.

This laboratory study covered only one corner of the overall gift purchase situation, but is consistent with the research cited earlier. From the perspective of a marketer of giftable products, there is an indication that for some products in some gift situations, the gift shopper may want to shop more rapidly than the own use buyer, so simplified shopping arrangements (such as a wedding gift book) could aid sales.

#### REFERENCES

- Russel W. Belk, "It's the Thought That Counts: A Signed Diagraph Analysis of Gift-Giving," The Journal of Consumer Research, Vol.3, No.3 (1976) 155-162.
- Carol A. Berning and Jacob Jacoby, "Patterns of Information Acquisition in New Product Purchases," The Journal of Consumer Research, Vol.1, No.2 (1974) 18-22.
- Roger M. Heeler and Chike Okechuku, "Brand Selection for Gift Giving Versus Personal Use," ASAC Conference (June, 1978) (Proceedings in press).
- Roger M. Heeler and Michael L. Ray, "Measure Validation in Marketing," Journal of Marketing Research, 9 (1972) 361-370.
- Jacob Jacoby, Robert W. Chestnut, Wayne D. Hoyer, David A. Sheluga, and Michael J. Donahue, "Psychometric Characteristics of Behavioural Process Data: Preliminary Findings on Validity and Reliability," Purdue Papers in Consumer Psychology, Paper No.172 (1977).
- 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," Purdue Papers in Consumer Psychology, Paper No.151 (1975) 306-314.
- C. Levi-Strauss, "The Principle of Reciprocity," in Lewis A. Coser and Bernard Rosenberg (eds.), Sociological Theory, (New York: MacMillan Co., 1965) 76.
- Adrian B. Ryans, "Consumer Gift Buying Behaviour: An Exploratory Analysis," American Marketing Association, Fall Conference (1977) (Proceedings in print).
- Barry Schwartz, "The Social Psychology of the Gift," The American Journal of Sociology, Vol.73, No.1 (1967) 1-11.
- Karl C. Weigl, "Perceived Risk and Information Search in a Gift-Buying Situation," Unpublished Master's Thesis, Purdue University (1975).



## CONSUMER GIFT-GIVING: OPENING THE BLACK BOX

Richard J. Lutz, University of California, Los Angeles

### Abstract

The three studies in this session on consumer gift-giving are characterized by a "black box" orientation and fail to reveal underlying determinants of gift-giving behavior. Future research needs to more systematically address the issue of motivations for gift-giving behavior.

### Introduction

Today we have heard the results of three studies pertaining to gift-giving behavior. Two of these studies were empirical in nature, while the third was essentially a literature review leading to a conceptual model of the gift-giving phenomenon. The question that must be addressed after hearing these three papers is the following: What do we know now about gift-giving that we didn't know prior to hearing these papers? The remainder of this discussion will attempt to answer that question.

### Gift Search Behavior

As noted by Banks in her paper, the search process for gifts has not been the subject of much past research. Thus, it is laudable that both the empirical papers presented today dealt with consumer search effort in the purchase of a gift versus search effort expended when purchasing for oneself. While the two studies are not directly comparable in all the issues they addressed, both studies did examine the amount of consumer effort expended in purchasing a "personal gift for a close friend" (Heeler, et al.); the gift situation reported on by Clarke and Belk was "a gift for a good friend." In each case, the gift purchase situation was contrasted with the purchase of the same item for "personal use."

Heeler, et al. concluded that there was no difference in search effort between the two situations, while Clarke and Belk observed a significant main effect, such that "effort expenditures for gifts [were] reported as higher than expenditures for personal product use." How can these apparently conflicting findings be resolved? In order to answer that question, we must examine in detail the studies which led to those conclusions.

### Recap of the Heeler, Francis, Okechuko and Reid Study

The study reported by Heeler, et al. really consists of two separate two-group experiments. In the first experiment, which will not be discussed at length here, students were asked to select information about a blender from an information display board (IDB), under one of two conditions--for their own use, or for a wedding gift. The rationale was that the wedding gift occasion constituted a "distant" gift-giving situation, although this assumption was never verified. In the second experiment, which is of more direct relevance to the present discussion, students again used an IDB to examine information regarding watches either as a gift for a close friend or for their personal use. This "close friend" manipulation was intended to represent a "close" gift-giving situation which again was an untested assumption. In the absence of manipulation checks, we have no evidence to support the contention that the two experiments in truth examined differences in two gift-giving situations differing in "proximity."

Subjects in Heeler, et al.'s. "close friend" vs. "personal use" experiment were male and female students at a "Polytechnic Institute." They were shown an IDB displaying 16 makes of watches with 10 attributes per watch. Eight of the watches were men's and eight were women's models, so that each subject presumably selected a friend from whichever sex he/she preferred in the "close friend" condition. We are given no information as to what proportion selected men's vs. women's watches. In any case, the effective size of the IDB was 10 attributes by eight brands for any given subject, regardless of sex or the sex of the imaginary recipient of the gift. Subjects selected information cell by cell until they reached a choice of watch.

Four dependent measures relevant to consumer search effort were recorded by an observer: (1) number of cells accessed, (2) search time, (3) number of brands consulted, and (4) number of attributes consulted. The former two measures showed no significant differences between the two cells. The latter two measures moved in opposite directions, but no tests of significance were reported for either measure. Based on these results, Heeler, et al. concluded that there were no differences in search effort between the "personal use" and "close friend" situations.

### Analysis of the Heeler, et al. Study

Internal validity. As noted above, there was no evidence presented with respect to the "proximity" manipulation. Similarly, there was no evidence to suggest that subjects in the two conditions of the watch experiment clearly perceived themselves to be in truly different situations. While the instructions have face validity (i.e., personal use vs. gift for close friend) one wonders just how much "mundane realism" the experimental conditions contained. Subjects were presented with a highly artificial task with no real expectations that either they or their "close friends" would actually receive the item in question. Thus, the two conditions may have, in fact, simply been viewed as a "game" in which the subject was trying to be a good player. In the absence of any realistic expectations regarding the consequences of their choices, would one really expect the subjects in the two conditions to behave differently?

External validity. The population being studied was quite narrow, and one can question whether any of the subjects in the experiment had ever previously considered purchasing a watch for a friend. Also in question, however, is the generalizability of the findings to the immediate sample employed. Only one product out of a plethora of possible gift items was studied; would the results be the same for all other products? Similarly, there are many conceivable situations in which one might consider giving a gift to a close friend. Would the results of the present study remain consistent across a variety of gift-giving occasions? In order to consider these questions, one should sample systematically from the domain of gift items and the domain of gift-giving situations.

Measurement reliability and validity. Little is known about the reliability of IDB measures of search behavior. Would the same results be obtained if the exact same people were subjected to the experiment a second time? With respect to validity, it is unclear that an IDB can

be used to address the question of search effort. By their very nature IDB's display information in such a way that subjects can obtain information with very little expenditure of effort. In contrast, actual shopping behavior is much more effortful in terms of time and inconvenience.

Thus, the true validity of IDB-derived measures is very much in question. The argument that only relative levels of effort were being examined rather than absolute levels of search pales in the face of possible test-treatment interaction effects. Consider a 2x2 experimental design in which the "treatment" factor is "personal use" versus "close friend," and the "test" factor is the shopping environment in which the treatment is administered, such that one level is an IDB and the other level is "the real world." It seems quite likely that different patterns of search would occur between the two test environments.

In sum, the evidence pertaining to internal validity, external validity, and measurement reliability and validity is quite weak for the Heeler, et al. study. This calls into question their basic conclusion with respect to differential search effort. Let us now turn to the Clarke and Belk study.

#### Recap of the Clarke and Belk Study

Clarke and Belk ostensibly used a gift-giving versus personal use manipulation to study the effects of "task involvement" on consumer research effort. They also used four different product categories to vary "product involvement." Since neither manipulation was checked within the study, we can safely ignore the involvement issue and focus on gift-giving.

Subjects in the Clarke and Belk study were undergraduate female students presented with two experimental conditions, in random order. In each condition, they were asked to respond to a series of intentions measures with respect to the (1) amount of time, (2) number of stores, and (3) amount of money to be spent in making a purchase decision. Although the exact measures used were never described, two forms of each measure were used, a "real" measure and a "relative" measure. Thus, a total of six dependent measures were obtained from each subject in each of two experimental conditions. In one condition, the subjects were asked to respond as if they were buying for personal use; in the other condition they responded as though they were buying a "gift for a good friend." The four product categories, which were completely crossed with the "involvement" manipulation, were bubble bath, blankets, jeans and record albums. The former two were chosen to represent low product involvement, while the latter two presumably were higher in product involvement.

Based upon a series of six univariate 2x4 factorial ANOVA, Clarke and Belk observed a general pattern of results in support of the hypothesis that high "task involvement" (i.e., gift) leads to more search effort than does low "task involvement" (i.e., personal use).

#### Analysis of the Clarke and Belk Study

Internal validity. Ignoring the issue of involvement, which was mentioned above, does the Clarke and Belk study effectively manipulate gift-giving versus personal use? While no direct manipulation checks were reported, the fact that rather large and systematic differences

were observed in the dependent variables across the two conditions leads one to believe that something was manipulated. The instructions were not fully presented so there is no basis for determining their "strength;" however, they do have face validity. It should be noted that Heeler, et al. used quite similar instructions; the difference here is that the experimental task of filling out a questionnaire is not as exotic as that of playing with an IDB. Thus, a very similar manipulation may have worked better for Clarke and Belk because the task was not so powerful as to overwhelm it. However, the manipulation is still very sterile and lacking in mundane realism. In particular, one wonders how often products like bubble bath, blankets, and even jeans are given as gifts by college coeds. Only record albums would seem to be reasonably common as gift items out of the four product categories studied. To the extent that the products are not ones commonly thought of in terms of gift-giving, motivations other than that implied by the manipulation may be entering into subjects' responses. Additionally subjects may find it difficult to provide adequate estimates of such behaviors.

External validity. As in the Heeler, et al. study, a very select subject population was investigated. To what extent are Clarke and Belk's results generalizable to their own subjects? With respect to sampling the product domain, Clarke and Belk used four products rather than only one. This is an improvement in quantity, but not necessarily in quality if the products are not representative of gift items typically selected by the subjects. Similar to Heeler, et al., Clarke and Belk did not specify either the person or the gift occasion, leaving that free to vary across subjects. Thus, the generalizability of the observed pattern over gift situations or recipients has not been established.

Measurement reliability and validity. Clarke and Belk's measures were not described in enough detail to understand them fully. All six appear to be assessed via single-item scales, leaving serious questions regarding their reliability and validity. In particular, the two "money" scales are questionable even in terms of their face validity as measures of search effort. It should also be noted that the six measures were probably fairly highly correlated and could possibly have been used to construct a single, more reliable measure of search effort.

In summary, examination of the Clarke and Belk study leaves little encouragement that their conclusion is generalizable to other people, other products, or other gift occasions.

#### Summary of the Two Empirical Studies

From the above description and analysis, it is impossible to resolve the conflict between the Heeler, et al. and Clarke and Belk studies. The two experiments used virtually identical instructions with respect to "personal use" versus "gift for friend," but the task environments and dependent measures were quite different. Neither study demonstrated adequate internal or external validity, nor were the measures utilized satisfactorily documented. In short, the results of neither study should in any way color subsequent investigators' thinking about the substantive phenomena of gift-giving behavior.

Of course, it is unrealistic to expect any single study to adequately sample person, product, recipient, and occasion domains. But it is the responsibility of researchers in the gift-giving arena to recognize these domains and their potential impacts on their studies' findings. Only when these domains have been

<sup>1</sup> It appears that an incorrect statistical model was used in the analysis, since subjects were not nested within cells.

adequately sampled will we have good descriptive information upon which to build a basis for truly understanding gift-giving behavior.

#### A Paradigm for Gift-Giving Research

Banks, in her paper, begins to move away from purely descriptive studies of gift-giving behavior and into the domain of explanatory constructs. The notions of reciprocity, identity formation, and interaction all seem to hold promise for future research in this area. However, none of these explanations appears in the paradigm presented later in the paper. In fact, the paradigm as presented seems devoid of any motivational constructs which might explain why the giver is buying the gift.

It is not entirely clear that consumer gift-giving behavior must be studied from the perspective of a paradigm different from that applied to other forms of consumer behavior. For instance, a general expectancy-times-value model should capture significant motives influencing gift-giving. None of the studies reported here today address the underlying determinants of gift-giving. For marketers to engage in actions designed to modify consumer gift-giving behaviors, a thorough understanding of these behaviors is necessary. Such an understanding must incorporate influences due to the type of gift occasion, the nature of the recipient, and intra-individual variables characterizing the giver.

Both of the empirical studies reported here operated at the level of brand choice. It would appear that two decisions more fundamental to the gift-giving domain are (1) the budget decision (i.e., how much to spend), and (2) product class selection. In fact, the gift occasion may be an extremely complex one in the sense of delineating product/market boundaries. For example, if the giver's primary requirement is that the gift in some way be unique, then the uniqueness attribute may be housed in an incredible number of alternative product forms. This goes back to the necessity for understanding giver motivations, for only then can the range of possible choice alternatives be identified.

#### Conclusion

In conclusion, the three papers in this session have added somewhat to our knowledge of gift-giving in a descriptive sense. Further, they have pointed out some directions for future gift-giving research of a more explanatory nature. Hopefully future research will pay more attention to motivational constructs and help to reveal the contents of the black box of gift-giving behavior.

Introduction

Is Gift-Giving A Legitimate Research Area?

Notwithstanding the fact that gift-giving has received little attention by the marketing profession, the study of gift-giving has a long history, primarily in the area of cultural anthropology. Belshaw's book (1965), "Traditional Exchange and Modern Markets," which is given short shrift -- and by only one of the three papers -- provides an exciting treatise on gift giving as both a medium of exchange and an instrument of social reciprocity. The studies of the Brobriands of Papua in New Guinea and the Kwakiutl Indians of northern British Columbia clearly point to the important role of gift giving both as a generator of power and as a medium for social exchange and reciprocity. Interwoven throughout these studies is the role of gifts in establishing ceremonial wealth and status.

From the viewpoint of marketing, there are a number of areas that are worthy of intensive study. They fall into distinct classifications:

- i) Product boundaries within the gift "situations." For example, little research has been done to establish the competitive set of products that enters the cognitive set for various gift-giving situations (family member, intimate friend, casual acquaintance, etc.). Along a similar vein, little research has been done on price-point boundaries for various gift-giving situations.
- ii) Product-line policy for manufacturers and retailers who depend heavily on gift-giving for a large proportion of their total sales. A sub-category of the above might look only at packaging issues for products that are purchased for self versus gifts.
- iii) Retail exploitation for specialty stores that focus on gifts (for example, Bowerings of Canada).
- iv) Information processing and buyer behavior in the gift purchasing arena.
- v) Gift-giving and reciprocity

All five areas above could contribute useful knowledge as an aid in the development of marketing strategy for either retailers or manufacturers. There may well be other areas worthy of study from a social or anthropological viewpoint but they are of less interest to marketers. Where do the three papers, presented here today, fit into the above research set?

Gift Giving: A Review And Interactive Paradigm

Ms. Bank's paper would appear to be the literature survey chapter of a dissertation and is therefore an appropriate starting point. While this discussant is unfamiliar with most of the gift-giving literature, this paper does have a solid core of literature which overlaps with the other two papers, particularly on the key studies by Belk, Levi-Strauss, Ryans, Schwartz, and Weigl. (See presentation papers for references.)

If the literature survey in this paper is fairly exhaustive, then it is clear that the research focus in the

past has been predominantly in classes iv) and v) from above, i.e., oriented towards social exchange theory, group and interactive boundaries, and identity formation. Within the marketing framework, some work has been done on information processing and risk reduction/risk perception. In addition, the relationship between perceived risk (social and financial) and the amount and type of information search has been examined. Some work has been done on frequency of gift giving, relevant gift categories for gift occasions and gift giving motivations/occasions. One of the key differentiating aspects of these studies is whether the gift itself is the focus of the study (the primary variable) or whether gift-giving is treated as a situational variable in order to study some aspect of the gift giving process. Few of the studies reported in this paper lead to significant marketing strategy implications. Yet one fact is inescapable. If the buyer behavior process in gift purchasing is significantly different from the buyer behavior process for "own purchase," or non-gift purchase, then the process itself is worthy of considerably more research effort as an aid in fine-tuning marketing strategy for both retailers and manufacturers who depend on the gift market for their survival.

The major contribution of the Banks paper is the presentation of the interactive gift-giving paradigm which dichotomizes the giver and receiver and provides a structure for research as one moves through the stages of gift giving. From a marketing perspective, only the first stage is really critical in terms of offering a fruitful research arena. That stage (the purchase stage) would include all the suggested research topics outlined earlier in this paper. As Banks concludes, little work has been done on relevant product sets, store sets, brands, product attributes, pricing strategy, packaging, etc.

Finally, the paucity of suggestions for research topics on the latter stages of the gift giving process confirms that only stage 1 of the paradigm is likely to receive future research attention.

The non-theoretical nature of the paradigm would introduce some problems for empirical testing. It is essentially descriptive in nature and not tied to any basic discipline. It does, however, provide a framework for analysis. The next stage would be a fuller exposition of buyer behavior theories in gift-giving/purchasing with the development of the relevant hypotheses.

Gift Versus Personal Use Brand Selection

This paper reports on an experiment, utilizing the information display board (IDB) methodology to test the level of "effort" extended in purchasing gifts for close versus distant friends. The experimental design involved the specification that blenders were appropriate for a "friend's wedding" while watches were appropriate as a gift for a "close" friend. Those decisions seem somewhat arbitrary and many have in fact, mitigated the findings. The experimental design is thus affected by both use occasion and by product. In addition, the (IDB) utilized 80 cells for watches and 160 cells for blenders. Thus, one might hypothesize in advance that information search for blenders would involve significantly more search (time, brands, attributes) than the search for watches, perhaps not 50 percent more, but some hypothe-

sized amount.

Data were collected on number of cells accessed, search time, number of brands consulted, number of attributes consulted, etc.

The analysis problem is apparent in Table 1 of the authors' paper. The four cells involved four exclusive groups who worked through the IDB on only one shopping trip; i) purchase blender for self; ii) purchase blender for wedding gift; iii) purchase watch for self and iv) purchase watch for close friend. The comparison of any two cells is impacted by differences in IDB cell numbers as well as occasion and recipient.

The search for blenders, for example, resulted in more information cells being accessed for both experimental groups compared to both groups who worked on the watches. While one might be prepared to conclude that buying blenders for self involves more search than buying blenders for a wedding gift, one cannot really say anything about blenders versus watches. This problem leads to non-comparability of data in Table 2 between watches and blenders.

It is also not clear why data on average prices paid is a good measure of "similar brand selection". Very different brand mixes could lead to the same average price paid by cell. Table 2 provides no data on statistical significance in mean scores or percentages.

Without a careful understanding of the methodological peculiarities, one might be tempted to conclude from the tables that blenders involve more search time than watches. That would be an erroneous conclusion. Without significance tests, one cannot reach any conclusions about the original hypotheses either. Finally, the authors' conclusion about marketing implications are fairly weak. Yet there are some interesting results worthy of note. First, information that was accessed fairly frequently for both watches and blenders included brand name and price. If additional research across other gift items provided the same results, then the implications for marketing strategy for both manufacturers and retailers are quite exciting. Catalogue showrooms, for example, market primarily on the basis of national brands at the lowest prices in the market. One might hypothesize that such outlets will be extremely successful, particularly since they appeal heavily to the gift-buying public. Second, each product has one critical specific attribute that seems to be very important; number of speeds for blenders and design for watches. The finding suggests a particular marketing focus for both products both in terms of advertising and package information/labelling.

#### The Effects Of Product Involvement And Task Definition On Anticipated Consumer Effort

Gift-giving is not the major focus of this paper. Rather, the research design examines product and situation involvement in purchasing behavior as the impact the search process in terms of stores shopped, time spent and amount of money spent under various conditions of high and low task importance and product involvement.

In the authors' view, the level of effort expended in searching for a product is related to both the degree of product involvement and the level of task involvement (buying for a gift is high task involvement). Their review of the relevant literature on both product and task involvement supports that hypothesis.

One of the key research issues is the extent to which three different measures of effort (stores shopped, time spent and money spent) are independent, interactive, or substitutes.

Task involvement was dichotomized as purchasing for self versus as a gift for a close friend. Product involvement was dichotomized into low (blanket, bubble bath) versus high (jeans and record album) but the rationale rests in an unpublished paper and is not defined for this study. The dependent variables are the three "effort" measures of stores shopped, time spent and money spent but they are projective rather than self-reports of past behavior. The reliability of projective measures is unknown.

It would have been extremely helpful if the authors had reported simple mean scores on the dependent variables by treatment cell. It is also not clear that the ANOVA analysis is the correct procedure. The real sample size is not 112 but the number of cells for which we have mean scores on the dependent variables (i.e., 4)...Figure 1 requires additional explanatory notes for interpretation. The tests for main effects on low versus high product involvement and low versus high task involvement would be correct.

There is no other research with which this author is familiar that supports the notion that price paid is a measure of shopping effort.

Given some of the weaknesses in the experimental design, the authors conclude that:

- i) high task involvement leads to high levels of effort on all three effort measurements,
- ii) higher product involvement sometimes leads to greater effort,
- iii) there is an interaction effect between task involvement and product involvement. Task involvement leads to higher anticipated effort for low but not high involvement products.

In marketing jargon, the authors conclude that higher task involvement (gift giving) can move a low involvement product to a higher level of involvement and therefore a higher level of effort. Thus, during peak purchase periods (e.g., Christmas), even products such as blankets and bubble bath, which have very different price levels, may become high involvement product classes. The problem is with the "level of involvement". It is never specified in the paper and there may well be thresholds below which no marketing implications would be forthcoming.

#### Directions For Future Research

From a marketing perspective, only in the case where meaningful differences (translatable to marketing strategy options) occur in the buyer behavior process for gift purchasing versus purchasing for non-gift purposes, would the research direction exemplified by these two papers be worth pursuing. In this author's opinion, neither paper generates very exciting findings from that perspective. The Heeler paper clearly shows that the types of information sought by gift and non-gift buyers is virtually identical. The Clarke paper indicates that manufacturers might want to offer a higher price range for products positioned in the gift market. They already do. They also look for wide distribution because gift purchasers shop more stores.

From the viewpoint of the manufacturers who are dependent on the gift market, there are much more interesting research questions. Is their product clearly perceived to be in the gift set and if not how do they get it there? How do they enhance the status of their product as an appropriate gift. What are the dimensions of "gifts"? What are the product boundaries and do they differ by type of gift (important versus non-important gift)? What are the relevant price ranges? What is the

most effective retail strategy? How can more gift-giving occasions be generated? How can gift giving be made more important. Can the traditional social reciprocity mores of older cultures be re-introduced and fostered? Can the notion of reciprocity be used more effectively in advertising copy? In short, can the mass market be manipulated into more expensive gift-giving. And how? Finally, how does the manufacturer generate a higher level of product involvement for his products?

## VALUE STRUCTURES AND CONSUMER BEHAVIOR

Jonathan Gutman, University of Southern California  
Donald E. Vinson, University of Southern California

### Abstract

This paper presents the results of an empirical investigation of the structure of human values. In addition to positing the relationship between values and behavior, the paper demonstrates a unique approach for exploring the tradeoffs consumers make in striving to achieve their values in a complex world.

### Introduction

It is difficult to imagine a single example of consumer behavior without reference to personal values. Even though the topic has been largely ignored by consumer researchers, the most cursory review of the social science literature demonstrates the important role of values in the study of both individual and collective behavior. In a recent interdisciplinary analysis of personal values, Clawson and Vinson (1977) suggested that:

The study of consumer values shows many signs of becoming a challenging area for research . . . values may prove to be one of the most 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.

While we have barely begun to scratch the surface of the subject that Toffler (1969) terms "so quick-silvery and complex," most serious students of human behavior appear to agree--values count!

### Personal Values

In the area of marketing and consumer behavior, value research has been heavily influenced by the theoretical and operational contributions of Milton Rokeach. "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). An individual's values represent a relatively small number of very centrally held evaluative beliefs. They are the ". . . cognitive representations and transformations of needs" and provide the criteria or standards by which judgments are made (Rokeach, 1973). Our values then, are responsible for the selection and maintenance of the ends or goals toward which we strive and, at the same time, regulate the manner in which this striving takes place.

Employing the Rokeach paradigm, consumer researchers have investigated differential product preferences (Vinson, Scott and Lamont, 1976), cross-cultured consumption patterns (Kanter, 1977; Munson, 1977), market segmentation potential (Vinson and Munson, 1975), consumer dissatisfaction (Scott and Lamont, 1973; Vinson, 1976), life style (Carman, 1977), and cognitive structure (Vinson and Nakanishi, 1977). In almost all of these cases, the data collection vehicle represented the Value Survey containing Rokeach's original 36 terminal and instrumental values.

### Value System Structure

Rokeach, as well as a large number of other value analysts, contend that values exist in a hierarchical, interconnected structure. That is, while all values are important and linked together, some values are more important than others. This notion is intuitively appealing for the average 22-year-old college student (the value "An Exciting Life" is usually rated more important than the value "Salvation"). Vinson, Scott, and Lamont (1976) suggest that the value structure exists as a central-peripheral dimension ranging from the most centrally held to the least centrally held. However, when college students are asked to rank-order the Value Survey, many complain that certain values "clump together" and these value "clumps" take on differential importance. "It is the rare and limiting case," Williams suggests (1968), that ". . . a person's behavior is guided over a considerable period of time by one and only one value . . . more often particular acts or sequences of acts are steered by multiple and changing clusters of values."

Apparently then, the conception of values being held independently of other values, rank-ordered from most centrally held to least centrally held is an inadequate representation of the consumer's value system. A more adequate system would be one in which some values are consonant with one another to the effect that the same behaviors are instrumental to achieving them. Some values by the same virtue, are contrary to one another in that they are mutually exclusive. Striving for one value or set of values means moving away from other values or sets of values. And, some values are indeed independent.

A major objective of this paper is to explore this proposition.

### Values and Behavior

In addition to not fully understanding the structure of the value system, consumer researchers have been rather obtuse in their explanation as to exactly how values are related to overt behavior. Not unlike other authors writing in the area of consumer behavior, Markin (1974) addresses this issue by stating, "A complete treatment, one that would trace the role of consumers' core values and their impact on such behavior as buying and consumption, is beyond the scope of our introductory treatment here." Little, if any, additional insight is provided by more recent researchers interested in the topic of personal values (Clawson and Vinson, 1977).

As end-states which give expression to human needs, we believe that values provide the goals toward which behavior is motivated. The immediate functions of values and value systems are to guide human action in daily situations. If we behave in all the ways prescribed by our values, we will be rewarded with all the end-states specified by these values (Rokeach, 1973, p. 14). In this sense, behavior is instrumental to the achievement of a particular value state. Behavior instrumentality implies that a specific behavior will either enhance or block the attainment of some value or constellation of values. Consumers are motivated to

engage in behaviors which will enhance the achievement of certain values; similarly, they are motivated to avoid those behaviors which are perceived to block the attainment of certain value states.

The research reported in this paper demonstrates the instrumentality of a variety of different behavior for a given set of personal values.

#### Measuring Values

As originally developed by Rokeach, the Value Survey requires subjects to rank-order the 36 terminal and instrumental values printed on gummed labels. For the purposes of consumer research, there are a number of problems associated with this technique. First, the ranking procedure forces the respondent to indicate differences where none may actually exist; equally attractive values are forced into separate rankings. In addition, wide gaps in preference are treated as no different from very small gaps. Second, many subjects verbalize Miller's admonition (1956) that most people cannot adequately evaluate more than a few items (7 plus or minus 2) at a time. And finally, the ranking instructions bias the rankings in favor of deprived values and against satiated values. "For instance, the lowest income respondents rate 'A Comfortable Life' and 'Clean' relatively high, while wealthy respondents rate them quite low" (Clawson and Vinson, 1977).

To avoid these problems, the vast majority of the empirical research previously cited employed the direct rating of each value. While this offers a number of methodological advantages, it doesn't directly address the issue of clusters of values that are consonant with one another or those which are mutually exclusive.

An important contribution of this paper is to introduce a new technique for exploring the relationships among human values. The results of this approach would enable researchers to explore the tradeoffs consumers must make in striving to achieve their values in a complex world. Specifically, the purpose of this research is to demonstrate these three types of relations (consonant, contrary, and independent) among values and to use these relationships to generate a values map in which the instrumentality of various behavior can be explored.

#### Methodology

To satisfy the objective of exploring the structure of personal values, data were collected by a unique form of paired-comparisons to generate a perceptual map. The stimuli to be mapped were values selected from Rokeach's terminal values (see Table 1). Because of the large number of items to be evaluated, a subset of 14 of the 18 terminal values was selected to reduce respondent fatigue. As can be seen from inspection of the values in Table 1, many of them are rather abstract with respect to their reference to end-states of existence. Also, these values can be satisfied in different ways by different people. At this stage of the research on values, it was decided to use the values as they were rather than to try to make them more specific. Any attempt at concretizing values would mainly serve to increase the idiosyncratic nature of the responses.

The scale used to elicit distances between the values was as follows:

Both value states can be achieved through the same behavior	1 2 3 4 5 6 7	The two value states are opposites--striving for one means moving away from the other
---	---------------	---

The context of the scale is behavioral--in any situation your behavior is leading you somewhere. Can you approach any two of these value states simultaneously or are they in opposition to one another?

One property of this scale is different than the often used "similar-different" scales. If two stimuli are unrelated or not comparable, the similar-different scale results in a "different" judgment. If two values are unrelated and striving for one is unrelated to the other, the respondent is most likely to mark the midpoint of the "can achieve both--they are opposite" scale.

Table 1

#### Values Employed in the Analysis

A Comfortable Life (A Prosperous Life)  
Social Recognition (respect, admiration)  
Pleasure (an enjoyable, leisurely life)  
A Sense of Accomplishment (lasting contribution)  
Self-Respect (self-esteem)  
A World At Peace (free of war and conflict)  
Equality (brotherhood, equal opportunity for all)  
Security (taking care of loved ones)  
Happiness (contentedness)  
Mature Love (sexual and spiritual intimacy)  
A World of Beauty (beauty of nature and the arts)  
Inner Harmony (freedom from inner conflict)  
True Friendship (close companionship)  
Wisdom (a mature understanding of life)

#### Property Vectors

As an aid to understanding the resulting space, ten behaviors were rated for their instrumentality in achieving the value states (see Table 4). Each of the behaviors was rated on a seven point scale as to whether it would "totally block the attainment" to "totally enhance the attainment" of each of the values. The behaviors chosen by the experimenters were thought to be relevant to the life style of the respondents, and related to the set of values selected.

#### Type of Analysis

The direct dissimilarity measures obtained from the subjects were analyzed according to the INDSICAL model (Carroll and Chang, 1970) and solutions were computed in three, two and one dimension. The INDSICAL model was used to retain some of the individual differences, hopefully adding to the depth of the solution.

#### Subjects

Subjects were students at the University of Southern California School of Business. Twenty-nine undergraduate students each rated the 91 value state pairs (14 X 13/2) and rated the instrumentality of the ten behaviors for each of the values.

#### Results

A minimum of three dimensions was required to describe the obtained space. The percentage of variance accounted for by the INDSICAL model was 36 percent. The average correlation coefficient across subjects was .59. While this figure is not as high as desired, it was felt that the interpretability of the obtained space provided a number of interesting insights and may provide assistance to others investigating the relationship between personal values and behavior.



Value State Map

The coordinates for the 14 value states in three dimensions are shown in Table 2. These 14 value states are mapped in Figure 1.

Figure 1

Three Dimensional Values Space

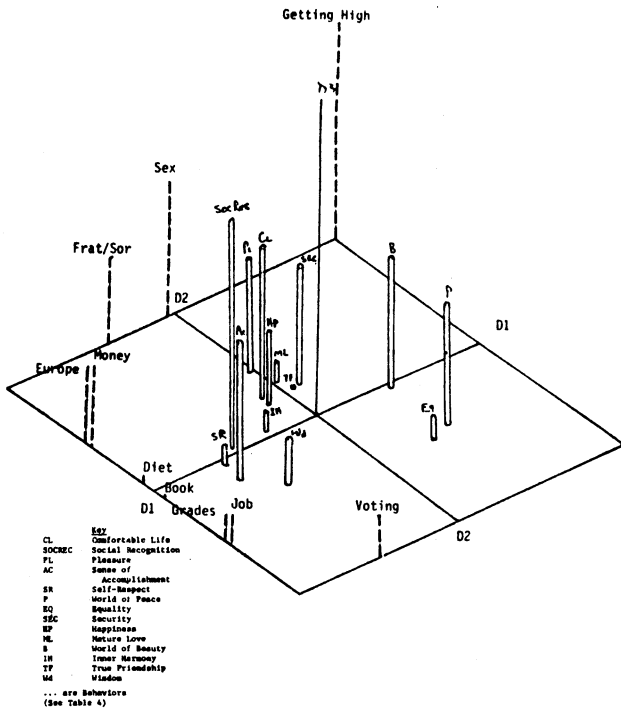


Table 2

Value Coordinates From INDSCAL Analysis In Three Dimensions

Value	Dimension		
	1	2	3
A Comfortable Life	-.036	.257	.640
Social Recognition	-.327	.037	.955
Pleasure	-.024	.429	.494
A Sense of Accomplishment	-.352	-.217	.560
Self-Respect	-.360	-.154	.099
A World of Peace	.467	-.388	.524
Equality	.392	-.450	.090
Security	.149	.168	.509
Happiness	-.049	.218	.312
Mature Love	.030	.293	.125
A World of Beauty	.385	-.025	.562
Inner Harmony	-.139	.007	.083
True Friendship	.112	.180	.000
Wisdom	-.242	-.356	.242

Interpretation of Dimensions

Certain features of the value state map are obvious at first glance. Beauty, Peace, and Equality are far removed from any of the other value states. Social Recognition bounds the space opposite these three values. Wisdom is not particularly near any of the other values. From a clustering perspective, it is unlikely that these values would be capable of achievement at the same time other values are being achieved.

Other value states are tightly clustered. True Friendship and Mature Love are close to one another and both are close to Happiness. Pleasure and A Comfortable Life are located near one another. Of particular interest is the fact that Happiness is located nearest

to the centroid of the space. From a personal perspective, this may be the most generalizable of the values. ("Whatever makes you happy...")

From a dimensional perspective, each of the three dimensions seem to be capable of an interpretation that increases our understanding of how people structure their values.

Dimension 1

The two poles of dimension 1 represent Personal Achievement versus Social Harmony (see Table 3). What seems to be reflected here is the fact that Social Recognition, Self-Respect, Accomplishment, and Wisdom are somewhat antithetical to Peace, Equality and a World of Beauty. This may be another way of saying that the pursuit of personal goals is contradictory with society's broader concerns. This may reflect the parochial, short time horizon of college students. It is also consistent with Rokeach's contention that terminal values are personal and social; that is, these values tend to be self-centered or society-centered, intrapersonal or interpersonal in focus (Rokeach, 1973, p. 8).

Dimension 2

Dimension 2 can be interpreted in terms of "Desires" versus "Oughts." In this context, Pleasure, Mature Love, Comfortable Life, and Security are opposed to Equality, Peace, and Wisdom. High ratings on the scale, which served as a basis for the distance measures input to the scaling algorithm, means these two sets of values cannot be achieved via the same behaviors.

Dimension 3

Dimension 3 represents one of the major aspects of college life (and to some respect, life in general)--social rewards and what it takes to get them as opposed to striving for some form of personal balance. Thus, we have Social Recognition opposed to Self-Respect, Inner Harmony, Mature Love, True Friendship, and Equality.

Table 3

Interpretation of Dimensions

Dimension 1	
Personal Achievement	Social Harmony
social recognition	peace
self-respect	equality
accomplishment	beauty
wisdom	
Dimension 2	
Desires	Oughts
pleasure	equality
mature love	peace
comfortable life	wisdom
security	
Dimension 3	
Social Rewards	Personal Balance
social recognition	self-respect
	inner harmony
	mature love
	true friendship
	equality

"The behaviors have been placed along the perimeter of the dimension 1-2 plane in Figure 1 such that a line between the point and the origin of the 3-space would define the plane. Table 4 shows the direction cosines for the behaviors which define their angular location in the space (Green and Rao, 1972, pp. 66-69). The calculations of the behaviors with the value states as mapped in the 3-space are very strong. With the exception of "reading a good book" ( $r = .507$ ) and "going to Europe" ( $r = .671$ ) all the correlations are greater than .76. Because rendering these angles and stimuli locations in three dimensions is somewhat precarious, Table 5 has been provided in which the instrumentality ratings for each behavior across the 14 values (along with the rank orders) are shown. Detailed inspection of Table 5 indicates, for example, that getting good grades and having a summer job are both instrumental in achieving Self-respect and A Sense of Accomplishment. Neither has much to do with A World of Peace or Equality. The point must be made that there is no prima facie reason why this should be so. Plenty of students can find a summer job helping underprivileged children or other activities that would help in achieving both of these sets of values. Our group of students is rather homogeneous (as will be discussed in connection with individual respondent saliences).

Table 4

Max "r" Property Fitting of Average Behavior Ratings in the Three Dimensional INDSICAL Stimulus Space

Behaviors	Direction Cosine on Dimension			Correlation Coefficient
	1	2	3	
1. Reading a good book	-.976	-.266	.002	.507
2. Sexual intercourse	-.040	.780	-.625	.857
3. Saving Money	-.826	.240	.510	.807
4. Getting High	.782	.612	-.122	.869
5. Matching one's diet	-.994	.124	.013	.762
6. Joining a fraternity or sorority	.402	.826	.396	.778
7. Getting a summer job	-.947	-.275	.170	.872
8. Studying to get good grades	-.957	-.253	.144	.937
9. Voting in a national election	-.344	.869	.356	.794
10. Going to Europe	.767	.233	.596	.671

Table 5

Average Instrumentality Ratings (Rankings) of Behaviors for Values

Values	Behaviors*									
	1	2	3	4	5	6	7	8	9	10
Comfortable Life	4.6(8)	5.2(6)	6.0(1)	3.6(3)	4.7(8)	4.5(7)	5.5(3)	5.4(6)	4.6(8)	5.7(5.5)
Social Recognition	4.5(9)	4.6(12)	5.5(5.5)	3.2(11)	5.1(5)	5.8(1)	5.1(7)	5.5(5)	4.9(5.5)	5.2(4)
Pleasure	5.8(2)	6.3(2)	5.2(7)	4.9(1)	4.8(7)	5.4(2)	4.7(9)	4.7(8)	4.2(8)	6.5(1)
Sense of Accomplishment	5.3(3)	4.9(10)	5.9(2)	2.6(14)	5.7(2)	4.4(8)	6.0(2)	6.4(1)	5.2(3)	5.8(3)
Self-respect	5.0(6)	5.2(7.5)	5.5(5.5)	2.7(13)	6.0(1)	4.6(6)	6.0(1)	6.2(2)	5.2(1.5)	5.1(3)
World of Peace	4.0(14)	4.2(14)	4.0(13)	3.9(6)	4.0(14)	4.1(12)	4.0(13)	4.0(13)	4.9(4)	4.2(12)
Equality	4.3(10)	4.9(11)	3.4(14)	4.1(5)	4.0(12)	3.5(14)	3.9(14)	3.9(14)	5.2(1.5)	2.9(13)
Security	4.1(12)	4.8(12)	5.8(3)	3.6(9)	4.6(10)	4.7(5)	4.7(6)	4.7(9)	4.9(5.5)	3.8(14)
Happiness	5.1(5)	6.2(3)	5.5(4)	4.2(3)	5.4(4)	4.9(4)	5.2(6)	5.7(7)	4.1(12)	6.0(2)
Future Love	4.2(12)	6.6(1)	2.1(10)	4.2(4)	5.0(6)	4.3(9)	4.1(11)	4.1(10)	4.0(14)	4.2(11)
World of Beauty	5.2(4)	5.1(9)	4.0(13)	4.2(2)	4.7(9)	4.1(13)	4.0(12)	3.9(13)	4.2(10.5)	5.7(7)
Inner Harmony	5.0(7)	5.2(7.5)	4.9(8)	3.5(10)	5.5(3)	3.4(13)	5.2(4.5)	5.5(4)	4.8(9)	4.2(9)
True Friendship	4.2(11)	5.8(4)	4.0(12)	3.8(7)	4.1(12)	5.2(3)	4.3(10)	4.0(11.5)	4.0(13)	4.5(10)
Wisdom	6.2(1)	5.7(5)	4.6(9)	2.9(12)	4.3(11)	4.3(10)	5.2(4.5)	5.6(2)	4.6(7)	5.7(5.5)

Further inspection shows the efficacy of joining a fraternity or sorority. This activity aids in achieving Social Recognition, Pleasure, True Friendship and Happiness. The interested reader can pursue any of the other behaviors to see what value states they aid. It is interesting, though to note that both voting and getting high are instrumental in achieving Equality.

Table 6 shows the individual saliences for the three dimensions. Although there are a few individuals who use one dimension much more than the other two dimensions, most subjects use three dimensions rather than one or two. There does not appear to be much of a basis for segmenting the group into subgroups with different value structures. This is partly a function of subject selection. Business students are not noted for being a highly differentiated group.

Table 6

Individual Subject Dimension Saliences from Aggregate Level INDSICAL Analysis of Direct Dissimilarities in Three Dimensions

Subject	Saliences on Dimension		
	1	2	3
1.	.22	.27	.29
2.	.48	.41	.23
3.	.18	.28	.64
4.	.04	.17	.01
5.	.46	.30	.30
6.	.37	.41	.30
7.	.27	.27	.21
8.	.60	.15	.27
9.	.52	.17	.40
10.	.28	.50	.07
11.	.21	.28	.31
12.	.46	.19	.12
13.	.34	.28	.19
14.	.32	.40	.26
15.	.19	.27	.57
16.	.22	.35	.01
17.	.64	.04	.23
18.	.34	.37	.31
19.	.55	.26	.09
20.	.34	.13	.43
21.	.59	.26	.09
22.	.53	.27	.18
23.	.29	.37	.21
24.	.15	.46	.18
25.	.21	.55	.37
26.	.63	.16	.20
27.	.20	.41	.35
28.	.32	.18	.44
29.	.29	.33	.43

Conclusions

This research has demonstrated Rokeach's terminal values are not all independent of one another. Some values are consonant and can be achieved via the same behaviors. Others are contrary to one another in that moving toward one means moving away from the other. There are several implications of this finding.

First, it suggests that a simple ranking of values for importance cannot do justice to their interconnected nature. At the very least, one would have to permit ties in the ordering. Rating the values for importance seems to be only a partial solution. While the rating procedure allows for equal importance rating, it does not force the respondent to deal with the reality that he or she cannot achieve all things--some value states can only be acquired at the expense of others.

Secondly, the results force us to think in terms of tradeoffs which have to be made in striving for valued states. If all values cannot be achieved simultaneously, values research has to address the same issues which confront attitude research. Indeed, we may have

come full-circle to Rosenberg's (1956) instrumentality-value analysis. Some values or clusters of values are blocked or attained by some products or behaviors.

Finally, the values-mapping procedure has provided a good model for understanding the alternatives facing our respondents: Personal Achievement vs. Social Harmony; Desires vs. Oughts; and Social Recognition vs. some form of balance in one's personal life. It seems likely that these dimensions are not independent of one another either. Further research will explore this issue by collecting "value-state as a source of satisfaction" data. We may then be able to develop a basis for segmentation by finding groups of consumers who are striving for the same value-locations in such dimensional space. Further research will also use products as property vectors in such spaces.

While the results of this paper are clearly exploratory in nature, we hope that they will contribute toward bringing value research more into the mainstream of consumer behavior research.

#### References

J. M. Carman, "Values and Consumption Patterns: A Closed Loop," Proceedings, Association for Consumer Research, 1978, 403-407.

J. D. Carroll 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.

C. J. Clawson and D. E. Vinson, "Human Values: A Historical and Interdisciplinary Analysis," Proceedings, Association for Consumer Research, 1978, 396-402.

P. E. Green and V. R. Rao, Applied Multidimensional Scaling: A Comparison of Approaches and Algorithms, (New York: Holt, Rinehart and Winston, Inc., 1972).

W. A. Henry, "Cultural Values Do Correlate With Consumer Behavior," Journal of Marketing Research, 13 (1976), 121-127.

D. L. Kanter, "The Europeanizing of America: A Study in Changing Values," Proceedings, Association for Consumer Research, 1978, 408-410.

R. J. Markin, Consumer Behavior: A Cognitive Orientation (New York: Macmillan, 1974).

G. A. Miller, "The Magical Number Seven, Plus-or-Minus Two: Some Limits on Our Capacity for Processing Information," Psychological Review, 63 (1956), 81-97.

J. M. Munson and S. H. McIntyre, "Personal Values: A Cross Cultural Assessment of Self Values and Values Attributed to a Distant Cultured Stereotype," Proceedings, Association for Consumer Research, 1978, 160-170.

Milton Rokeach, "A Theory of Organization and Change Within Value-Attitude Systems," Journal of Social Sciences, 24 (1968), 13-33.

Milton Rokeach, Beliefs, Attitudes, and Values, (San Francisco: Jossey-Bass, 1968).

Milton Rokeach, The Nature of Human Values, (New York: Macmillan, 1973).

M. J. Rosenberg, "Cognitive Structure and Attitudinal Effect," Journal of Abnormal and Social Psychology, 53 (1956), 367-372.

J. E. Scott and L. M. Lamont, "Relating Consumer Values to Consumer Research: A Model and Method for Investigation," in Thomas W. Greer, ed., Increasing Marketing Productivity (Chicago: American Marketing Association, 1973), 283-288.

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.

D. E. Vinson, "Personal Values as a Dimension of Consumer Discontent," in B. Greenberg and D. Bellinger, 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), 313-317.

D. E. Vinson and M. Nakanishi, "The Structural Composition of the Consumer's Value-Attitude System," Working Paper, College of Business, University of Southern California.

D. E. Vinson, J. E. Scott, and C. M. Lamont, "The Role of Personal Values in Marketing and Consumer Behavior," Journal of Marketing, (1977), 44-50.

R. M. Williams, International Encyclopedia of the Social Sciences, (New York: Macmillan, 1968), p. 287.

## MEASURING THE BASES OF SOCIAL POWER

John L. Swasy (student), University of California, Los Angeles

### Abstract

The French and Raven conceptualization of social power and previous operationalizations of the bases of social power are reviewed. The development of an instrument to measure perceived social power is presented.

### Introduction

The topic of social power is quite complex and has been described by many different sociological and psychological theories. Psychoanalytic theory, field theory, exchange theory, as well as the persuasion and attitude change literature all deal with aspects of the influence one person has on another. For the consumer behavior areas of family decision making and attitude change, interpersonal influence processes are of particular importance.

Because social power has been studied from so many perspectives, it's understandable that a wide variety of definitions and operationalizations have resulted. The terms "influence," "power," "decision making," "authority" and other terms have often been used interchangeably. In the family power structure literature, for example, the wide variety of definitions and measures has led some researchers to question the usefulness of the power concept (Safilios-Rothschild, 1970). A review of the social power literature indicates that one basic problem is the lack of a proven scale for measuring social power.

This study attempts to provide such a scale. Specifically it will review the French and Raven conceptualization of social power, previous operationalizations and then present the development of an instrument to operationalize social power.

### Theoretical Framework

Previous works have offered a variety of definitions of social power and influence (Pollard and Mitchell, 1972). A commonly cited definition is that of French and Raven (1959) and Cartwright and Zander (1968) who define influence as "a change in cognition, attitude, behavior or emotion of one person which can be attributed to another agent." Social power is "the potential influence of one person over another" (Cartwright and Zander, p. 316). Thus, power is the total possible change which one social agent can cause in another person's attitude, behavior, beliefs, etc.

This conceptualization follows from the early work of Kurt Lewin (1951), who posited a model of human behavior based upon forces within one's life space. In this model, behavior was explained as the resultant force caused by tensions and needs which originated within (or acted upon) the life space. The social influence of person A over person B was defined in terms of the force A could bring to bear on B and the resistance B could offer. Mathematically Lewin defined the Power of A over B as the quotient of the maximum force that A could (or possibly could) induce on B and the maximum resistance that B could offer. Cartwright (1959) reconceptualized power not as the quotient but as "the maximum strength of the resultant force which A can set in that direction at that time" (Cartwright, 1959, p. 193). Thus, according to this conceptualization, the extent to which agent A has control over B's behavior will depend on the magnitude of the force which he can bring to bear on B and

upon B's resistance.

While Lewin's and Cartwright's representation of social power and influence strongly mirror the physical sciences' notion of force vectors, they do include several important psychological aspects. First, for one person to "have power over B" does not necessarily mean that A is having an observable effect on B. If other situational and internal forces balance A's force, then no change will occur. Secondly, it is not necessary for A to act in order to have power over B. B's anticipation of A's reaction is sufficient for A to have power.

While the outcome of an influence exchange between two persons was very important to these early theorists, the process of influence and the reasons for social power were also described. Cartwright notes that two factors determine A's power over B. These are the resources of A and the needs or "motive base" of B. Thus, "an act of A must tap a motive base in order for it to activate a force." (Cartwright, 1959, p. 205) These motive bases were described by French and Raven (1959) and Raven (1965) as being composed of six different types--reward, coercive, referent, legitimate, expert and informational.

### Bases of Social Power

The Reward power of an influencer (A) over person B is based on the ability to mediate positive outcomes and to remove or decrease negative outcomes received by B. The strength of A's power is thus dependent on the size of the mediated outcome and the probability that A can mediate the reward. French and Raven also suggest that the amount of influence from this power base will depend on the subjective probability of reward for conformity, minus the subjective probability of obtaining the reward for nonconformity. In addition, with reward power, A's influence is usually limited to those domains where A has the ability to reward. Finally, since A's knowledge of B's compliance is necessary in order for B to be rewarded, B will want to relate his compliance to A and maintain A's surveillance of B's behavior. This surveillance by A is also important for another power base, coercion.

Coercive power rests on B's belief that A will punish him for not complying. The strength of the coercive power depends on the magnitude of punishment times the difference between the probability of punishment for nonconformity and the probability of punishment for conformity.

Often, reward and coercion are viewed as opposite ends of a one-dimensional concept. The primary distinction between these two power bases rests on the influencee's attraction to the influencer. With reward, an increase in attraction should result and lead to a higher dependency on the influencer. However, with coercive power the influencee's attraction toward the influencer should decrease and cause the influencee to avoid further encounters with the influencer.

The third basis for social power is referent power, which results largely from the influencee's feelings of identification with the influencer and desire to maintain similarity with the influencer.<sup>1</sup>

<sup>1</sup>Kelley's "comparison function" (1952) of a reference group is consistent with this power basis. However,

The fourth basis for social power, legitimate power, stems from internalized values of the influencee which dictate that the influencer has the right of influence and that the influencee is obligated to obey. These internalized values may reflect cultural values and norms, group norms, or role prescriptions. The influencee's underlying feelings of "oughtness" toward the compliant behavior and "has the right" toward the influencer characterize this power base.

Because this power base is often derived from many complex societal and personal values, the range of legitimate power A has over B will vary across situations. To the extent that A's legitimate power is derived from a general value or belief of B, A would be expected to have power across many situations. However, if the power rests on limited role characteristics, it is unlikely that A will have power outside the role situation.

One paradoxical type of legitimate influence is that of the "dependent" and "powerless." In some situations a person who has little power can emphasize his powerlessness and influence the more powerful. For example, a person may appear "helpless" in order to get another to perform a task. Schopler and Bateson (1965) suggest that the "power of dependence" is a form of legitimate power since it involves the acceptance of an influencee role relationship. Thus, the more powerful is obligated to perform the duties associated with the powerful role position and submit to the demands of the influencee.

The expert power basis stems from the influencee's attribution of superior skills or knowledge to the influencer. The degree of expert power is a function of the amount of knowledge the influencee has and the degree to which the knowledge or skills of the influencer are appropriate for a given situation.

The final power basis, information, differs from the previous bases in that it is "independent" of the influencing agent (Deutsch and Gerard, 1955). That is, this power stems from the "logic," "reasoning" or importance of the communication provided by the influencing agent and independent of the communicator. The content of the communication alone leads to changes in belief structures, behavior, attitudes, etc. In most situations it is difficult to independently distinguish expert and information social power. In fact Collins and Raven (1969) suggest that informational influence may follow only after some degree of expert power is perceived by the influencee (p. 184).

Although there has been considerable research on social power and the French and Raven bases, there has been little to improve the operationalizations of these concepts.

#### Operationalizations

One method of operationalizing social power is by creating an experimental treatment situation depicting the desired social power. The effect of the social power treatment on various dependent variables can then be tested. Busch and Wilson (1976) employed this method in a study on the effect of life insurance salesman's expert and referent bases of social power on subject's trust in the salesman, overall attitude toward life in-

the "normative function" of a reference group is analogous to the reward and coercion power bases discussed above. French and Raven note: "...to the extent that O (influencer) mediates the sanctions (i.e., has means of control over P (influencee) we are dealing with coercive and reward power, but to the extent that P avoids discomfort or gains satisfaction by conformity based on identification regardless of O's responses we are dealing with referent power." (French and Raven, 1959, p. 162).

surance and several behavioral measures. Low and high expertise power was created by presenting the salesman as either "above average" or "excellent" on six attributes of selling ability and training. The referent conditions of low and high similarity were created by using Byrne's attitude similarity procedure (Byrne, 1961, pp. 713-5). In this method, the influencer is presented as either being similar or dissimilar to the subject on a list of attitude topics.

Although the operationalizations employed in this study appeared at least on the surface, to be reasonable, the authors reported only marginal differences in referent power between the high referent-low expert and low referent-high expert conditions. This comparison was made using manipulation check measures for referent power which were two 7 point items ("how much they (subject) would like the salesman if he met him" and "would they enjoy working with him in a research experiment.")<sup>2</sup> Thus, while these items have reportedly high reliability, (0.85, Byrne, D. and D. Nelson, 1965), some questions arise regarding the validity of using measures of attraction as surrogates for referent power in a situation where several power bases are operant. From this one might suggest that either the concept of referent power, the method of operationalization, or the manipulation check measures be revised.

Another method of operationalizing social power in an experimental design is via a cartoon representation (Leet-Pellegrini and Rubin, 1974), (Raven, 1973). This consists of a series of cartoons illustrating an influence situation. In the Leet-Pellegrini and Rubin study, the first cartoon panel showed a young adult standing in front of a doorway at a city street corner. The second panel depicted a police officer saying to the young adult, "Would you please move away from here." The third panel to be selected operationalized different bases of power (e.g., (Reward) "If you do, I will keep it in mind--there are ways that I can help you," (Legitimate) "I'm asking you to do this as a part of my job.") While this cartoon method provides an alternative to the film presentation of Busch and Wilson, the need for manipulation checks remains. Such checks were not presented in the Leet-Pellegrini and Rubin study.

While a thorough review of common experimental treatment operationalizations of the other social power bases is beyond the scope of this paper, it should be noted that they seldom are more sophisticated than that of those mentioned above. Declarations by the experimenter regarding the influencer's ability to control monetary payments or electrical shocks, or attributed knowledge and experience to the influencer are also typical.

Another perspective on the measurement of social power is provided by Raven, Centers and Rodrigues (1975). In this study, the authors sought to determine the differential susceptibility to a social power base as utilized by one's spouse. Following the introductory statement: Now, there are many cases where your wife/husband asks you to do something and you do it, even though you may not see clearly why it should be done. We are interested in finding out why you might do as your wife/husband asks, so I will give you some possible reasons and would like you to tell me how likely each of these reasons is.

<sup>2</sup>The manipulation check for expert power was a one item, nine point scale--"I feel that the knowledge and competence of the salesman who is being studied here is"--with end points of "slightly below average" and "definitely above average." The reliability and validity of this item was not discussed.

The subjects rated the following statements on a three point scale (very unlikely, somewhat likely, very likely):

1. Because if you did so then she/he would do something nice for you in return,
2. Because if you did not do so, he/she might do or say something which would be unpleasant for you in return,
3. Because he/she knew what was best in this case and so I did what he/she asked you to do,
4. Because you felt that she/he had a right to ask you to do this and you felt obligated to do as she/he asked, and
5. Because you felt that given you were both part of the same family, you should see eye-to-eye on these matters.

While the face validity of these items appears to be high, the authors presented no information on the reliability and validity of these measures.

This brief review has illustrated the most common methods of operationalizing the French and Raven bases of social power and their shortcomings. The effectiveness of the experimental treatments is often not verified by manipulation check measures. Also, in cases where the manipulation check is not consistent with the treatment, one may question either the treatment or the manipulation check measure unless reliable and valid check measures are employed. The use of rating and ranking measures is also subject to deficiencies. Often, content and other types of validity for the items are assumed. In addition, single item measures are subject to many sources of error and do not allow for the assessment of reliability.

Thus while these approaches provide convenient operationalizations of social power, they do not answer issues of measurement reliability and validity. There is a need for more valid and reliable measures of social power.

#### Method

For this study a Likert approach was selected. Some of the assumptions which underlie the Likert procedure are: (1) the concept being measured is unidimensional; (2) the intervals between adjacent responses are equal; (3) the intervals are equal across items; and (4) a "positive" direction can be determined for each item (Runkle and McGrath, 1972, p. 314).

The first assumption will be addressed in the analysis section. The second and third assumptions are commonly made in Likert scaling situations and do not appear to be unreasonable for this study. The fourth assumption also seemed reasonable since trained judges were used to select the scale items.

#### Overview

A Likert scale development format was used for this study and included the following major steps: (1) A large pool of belief statements which reflect different characteristics of a social power situation was generated; (2) Judges rated these items to select those which clearly indicate a particular type of power; (3) Influence situations were generated and judged for use as standards for scale development. For each power basis, two situations were selected which depicted that social power base; (4) Subjects assumed the role of the influencee in a situation and responded to multiple items reflecting each basis of power; (5) Items reflecting each social power basis were analyzed to determine the most reliable items for the final scale; (6) Power base scores across situations were compared as a check on predictive validity.

#### Scale Items

The items generated by the author, depicted characteristics of either an influence situation in general, a particular relationship between the influencer and influencee, abilities or resources of the influencer, etc. One hundred and fifty items were generated, approximately 25 for each French and Raven power type.

While this process does not ensure a thorough sampling of the population of the social power characteristics for each power basis being measured, it is felt that the large number of initial items provides an acceptable representation of social power characteristics.

These items were given to six judges<sup>3</sup> familiar with the French and Raven typology. Each item was judged as to whether or not it was an indicator of each of the French and Raven types of power. A judge could designate the item to represent more than one power type.

An item was designated by the author as an acceptable indicator if:

1. A minimum of five of the six judges classified the item as an indicator for the same power type; and
2. The item also was not classified as an indicator in the other categories a total of 3 or more times (either within another category or over several categories.)

This procedure required reasonably high interjudge consistency, but still tolerated a small degree of inevitable variability present in judging tasks of psychological concepts. Eighty-five items were selected by the judges as acceptable indicators, 12-16 items for each type of power.

#### Scenarios

Scenarios were used to represent influence situations.<sup>4</sup> (See Figure 1 for examples.) In this case, the subjects were instructed to assume the role of the influencee in a given scenario, and then indicate their agreement or disagreement to each of the 85 belief statements selected by the judges.

An important aspect in the development of the scenarios was the need to generate situations which clearly depicted only one type of social power. This was necessary to ensure that the later scale results could be validated with a known situation treatment (i.e., the "expertise" power score would be highest in the expertise situations.) The second issue was the need to generate two situations depicting each type of power. This replication of situation treatments was intended to ensure that the scales developed from this research would have some degree of generalizability.

Situations were generated by the author and reviewed by the same panel of judges. The six judges rated thirty-six situations for each type of influence on a 1 to 5 scale having end points of "does not depict this type of power" and "clearly depicts this type of social power." For each type of power these ratings were summed across judges. Thus each situation had six scores (one for each type of social power) ranging from the lowest possible (6) to the highest (30). The final situations

<sup>3</sup> I wish to thank George Belch, James Bettman, Harold Kassarijian, Richard Lutz, Michael Munson, and Bertram Raven for their assistance and helpful comments.

<sup>4</sup> Because of the costs associated with using actual situations, a scenario approach was utilized. This approach allowed for the development of clear examples of each power basis. While using this approach lessens the external validity of the results, it does allow other researchers to easily and rather inexpensively test the findings.

were selected on the basis of being a clear representation of one social power type. For all the situations except for the expert situations, the selections were quite direct. For the expert situations, it was difficult to generate a situation free of informational power. The final expert situations used for this study were agreed upon by the judges to be clear examples of expertise, but also possibly examples of information power. This confounding is understandable considering the close theoretical meaning of these two types of social influence. Situation #4, an example of legitimate power, deserves comment. This situation depicts the "power of the powerless." In this type of situation persons who have few resources in the sense of power and influence, have a "legitimate" right to influence those who are more powerful. Thus either because of societal or personal norms, the more powerful person is obligated to conform to the request of the "powerless" (e.g., a prosperous person giving money to a beggar).

FIGURE 1  
SELECTED SITUATION SCENARIOS

(#2 EXPERT): You are the owner of a large manufacturing company. One of your customers has filed a lawsuit against the company. The customer's lawsuit is based on a very technical aspect of the law which you do not understand. You decide to contact an attorney (A) who specializes in this area of the law. After talking with the attorney (A) you are still uncertain about the technical aspects of the case. Even though you do not understand his reasoning you do as he says.  
 (#4 LEGITIMATE: (POWER OF THE POWERLESS) You are leaving your office building when a beggar (A) approaches asking for money. You give him money.  
 (#12 REWARD): You are talking with a car salesman (A). For 30 minutes you barter over features and price. Finally the salesman offers to include air-conditioning free if you buy. You accept.

Procedure

The final instrument consisted of one scenario and the 85 test items which were arranged in random order.<sup>5</sup>

Subjects for this research were undergraduate accounting majors at UCLA. A total of 321 were used in the final analysis. Approximately 25 other subjects were eliminated from the study because of incomplete questionnaires. These incompletes did not occur in any particular class or any one influence treatment. This relatively homogeneous population of subjects restricts the generalizability of the results.

Prior to the distribution of the test instrument the subjects were asked to participate in a social science research project and were told that their participation was completely voluntary. They were also informed that if they felt that they could not perform the task in a conscientious manner, to please return the questionnaire unmarked or incomplete. The required time for completion of this task ranged between 15 and 20 minutes.

Analysis

In Likert test development an index of the items' discriminating power serves as the evaluation criterion. Items with high discriminating power are selected for the final form. This scaling procedure implies a "single common factor" model (Green, 1954). A factor analytic scale analysis was chosen to construct the final

<sup>5</sup> Perhaps a better procedure would have been to word half the items in the negative to protect against possible response set biases. Since efforts were taken to encourage responsible participation by the subjects and no response patterns were observed in the data, it is felt that the procedure used is acceptable.

scales. (McKelvy, 1976). This approach is quite suitable for "subjective areas where external criterion variables have not yet been identified or where the emphasis is on discovering and defining new theoretical constructs and accompanying operational measures." (McKelvy, 1976, p. 2.) The lack of a clear external criterion is particularly characteristic of this research.

The factor analysis was performed for each set of items (basis of power) across all situations. In addition to providing the factor loadings for each item, this measurement set analysis program provides Cronbach's (1951) alpha coefficient for each set of items, starting with the two highest loading items, then the three highest, etc. This calculation was made for all sets of items having at least a commonality of 0.35 for the smallest loading item. While this cut-off is arbitrary, it appears to be reasonable for this study. The calculation of alpha allows for an assessment of the relationship between test length and internal consistency. While some research situations may require alternative criteria for item selection (i.e., a wide content of scale items), the issue of internal consistency is still important.

The major guidelines for using this program in a uni-factor solution include: Items should be written to measure a well-defined construct; the number of cases should be at least three times the number of items; and the data should be multivariate normal and homoscedastic. Except for the third assumption, these requirements were met. The effect of the departure from the third assumption is that the alpha coefficients are, at best, estimates.

Results

This section will first briefly review the overall "goodness of fit" for the uni-factor scales. Then each scale will be briefly reviewed. Table 1 presents the information on the factor analysis of each scale. The "total common variance with principal factor solution" indicates the degree of overlap between the test items, regardless of whether the 'overlap' occurs in a one, two, or n-dimensional space. Thus, these values suggest that even though the items were designed and rated by judges to be measuring the same concept, the "common" portion of the items account for at best 52.9% of the item scores' variance.

TABLE 1  
FACTOR ANALYSIS DATA FOR THE ITEM ANALYSIS

	Number of Variables	Total Common Variance with Principal Factor Solution (% of Total Variance)	Total Common Variance with Uni-Factor Solution (% of Total: Var. Common Var.)	
Coercion	12	50.6%	34.2	67.6
Expert	14	52.9	35.5	67.2
Legitimate	16	47.9	24.0	50.0
Reward	16	50.6	29.8	59.0
Informational	12	42.7	25.2	59.0
Referent	15	52.5	33.3	63.4

A "total common variance with the uni-factor solution" column in Table 1 indicates the degree to which one dimension accounts for between 24 and 35% of the total variance among the test items (e.g., a one dimensional scale accounts for about 25% of the "total" variance). Considering only the common variance, the proportion attributable to the single factor ranges between 50.0 and 67.6% (e.g., a one dimensional scale accounts for

about 50% of the "true" variance). These figures suggest that the assumption of a uni-dimensional concept is a reasonable one.

For the six social power (SP) scales discussed below, the factor loadings and alpha coefficients for items with communality greater than .35 are presented in Table 2.

While an appropriate reliability level depends on the variety in item content desired and the reliability of other constructs in a given research situation, a commonly accepted reliability level for most basic psychological research is 0.70 (see Nunnally, p. 226).

The coercion SP items, as expected, centered around potential harm and punishment from the influencer. The item loadings were consistently high and an alpha level of over 0.80 was obtained with the first three items.

For the expertise SP scale the item loadings were consistently high and the alpha values were equally acceptable. Trusting the influencer's knowledge and experience were the primary elements in this factor.

The legitimate SP scale items had fairly high factor loadings. However the best alpha level which could be obtained with items having an estimated communality greater than 0.35 was 0.59. One possible explanation for this is the limited range in item variances which resulted for some situations. Since the "power of the powerless" situation (#4) is a rather unique example of legitimate power, many of the legitimate SP scale items were inappropriate. Thus, having only one situation to represent the higher legitimate power value and 11 situations having small values could cause these lower intercorrelations. Since the same item generation and judging procedures were used for this SP scale as the others, it is believed that the items are indicative of legitimate power and that the lower alpha levels for this scale are primarily due to the limited array of situations used here.

The SP scale for reward is composed of items reflecting the non-attainment of rewards for non-compliance and the importance of the rewards as the motivation for compliant behavior. Loadings greater than 0.5 were obtained for most of these items. To achieve an alpha level of 0.8, five items were required.

As expected, the items dealing with the "logic," "good reasoning" and "sense" of the influencer's message are the best items for the information scale. While all the items have fairly high factor loadings, the best internal consistency (0.74) is achieved by the first three items.

Consistently high loadings were obtained for the referent SP scale also. Similarity in opinions and values and the desirability and goodness of being similar to the influencer were the principal identifiers of this scale. The first three items of this scale achieved an alpha of 0.80.

In summary, the scales for the French and Raven power bases provide reasonable levels of internal consistency. The lower reliability for the legitimate power scale appears to be the result of the situations used in this study. Since the procedures used in generating and judging the scale item were the same for all the scales, it is felt that the items for legitimate power will have acceptable internal consistency in most legitimate power situations.

TABLE 2  
SOCIAL POWER SCALES

Item	Loading	Alpha*
<b>REWARD</b>		
If I do not comply with A, I will not be rewarded.	.67	-
The only reason for doing as A suggests is to obtain good things in return.	.67	.67
I want to do as A suggests only because of the good things A will give me for complying.	.66	.74
A has the ability to reward me (in some manner) if I do as A suggests.	.65	.76
If I do not do as A suggests I will not receive good things from A.	.63	.81
In this situation I am dependent on A's willingness to grant me good things.	.60	.82
<b>REFERENCE</b>		
In general, A's opinions and values are similar to mine.	.77	-
Being similar to A is good.	.68	.71
I want to be similar to A.	.57	.80
In this situation my attitudes are similar to A's.	.65	.79
I would like to act very similar to the way A would act in this situation.	.63	.81
In this situation my behavior is similar to A's.	.62	.83
<b>INFORMATION</b>		
The information provided by A about this situation makes sense.	.67	-
The information A provided is logical.	.62	.67
I will seriously consider A's request because it is based on good reasoning.	.61	.74
<b>COERCION</b>		
A can harm me in some manner if I do not do as A suggests.	.77	-
If I do not do as A suggests, A will punish me.	.76	.77
Something bad will happen to me if I don't do as A requests and A finds out.	.72	.81
I had better do as A suggests in order to prevent something bad from happening to me.	.65	.82
A might do something which is unpleasant to those who do not do as A suggests.	.64	.84
<b>EXPERTISE</b>		
I trust A's judgment.	.74	-
A's expertise makes him/her more likely to be right.	.73	.63
A has a lot of experience and usually knows best.	.70	.74
A knows best in this situation.	.69	.80
A's knowledge usually makes him/her right.	.66	.83
I trust A's judgment in this situation.	.65	.85
In this situation I don't know as much about what should be done as A does.	.61	.85
A is intelligent.	.60	.86

(Cont'd. on next page)



TABLE 2 (Cont'd.)

LEGITIMATE			
It is my duty to comply with A.	.66	-	
Because of A's position he has the right to influence my behavior.	.61	.38	
I am obligated to do as A suggests.	.60	.59	

\*Alpha based on this item and all items with a high er loading. Items with estimated communalities of less than 0.35 are not entered into a calculation of alpha.

The predictive validity of the SP scales was examined by performing a one way ANOVA, with the situations as the treatments and a power score as the dependent variable. Thus, six ANOVAs were run one for each powerbase.

If the scales are actually measuring their respective power bases and the situations truly depict a given power base situation, then one would expect that significant differences in the power scores should occur across the situations. Also one would expect that the situations depicting a given social power base should have the highest mean score. Table 3 presents this analysis, using the three highest loading items in each scale to calculate the scores. Newman-Keuls tests of differences are also presented (  $P < .05$  ).

The significant F ratios and the higher mean scores for the appropriate situations suggest that the scales and situations were good operationalizations of the power constructs. For all six power bases, one of the appropriate situations had the highest mean score for the power scale. Except for the legitimate social power situation number 4 (the "power of the powerless"), the second situation for each power base was either second or third highest on the appropriate power score.

Discussion

The results of this study indicate that reliable measures of perceived social power can be obtained with relatively few items. Using the three highest loading items for each of the six French and Raven power bases results in a scale of reasonable total length and acceptable levels of internal consistency.

While it is desirable to have a scale with high internal consistency, the validity of the scale is equally important. A high level of content validity for this scale was sought by generating and selecting items which represented clear theoretical constructs. These constructs were the French and Raven bases of social power which have been widely cited and used in much of the social power literature. Also, the use of a panel of trained judges to screen the test items, and the "face" validity of the scale items suggest that the social power scales have an acceptable level of content validity.

For the assessment of predictive validity, criterion measures were created by generating scenarios which clearly depicted a given basis of social power. Using these "known" situations as standards, the predictive ability of the scales was examined. The consistency between the scale values and situations suggest that the scales have reasonable predictive validity.

The important aspect of construct validity remains to be tested. The replication of relationships between social power and other psychological constructs which were derived in laboratory experiments would serve as good indicators of construct validity.

An area of interest to attitude change researchers which might provide evidence for construct validity would be a

TABLE 3

ONE WAY ANOVA AND NEWMAN-KEULS TEST OF DIFFERENCES IN SOCIAL POWER SCORES ACROSS SITUATIONS

TYPE OF INFLUENCE EXPERT			REWARD			INFORMATION			REFERENT			LEGITIMATE			COERCION		
Situation Number	1st Mean <sup>*2</sup>	diff. sit.	Sit. No.	1st Mean	diff. sit.	Sit. No.	1st Mean	diff. sit.	Sit. No.	1st Mean	diff. sit.	Sit. No.	1st Mean	diff. sit.	Sit. No.	1st Mean	diff. sit.
7	6.34	5 <sup>*3</sup>	10	5.73	5	8	8.22	10	7	5.03	8	10	5.30	5	10	4.78	3
4	6.59	5	4	6.40	3	4	8.96	6	4	5.48	11	6	5.95	11	12	4.81	3
8	7.22	3	9	6.43	3	7	9.06	6	8	6.45	5	④	6.07	11	1	5.66	3
12	7.37	3	6	6.82	12	11	9.66	5	11	7.53	10	9	6.33	11	9	5.66	3
11	7.80	10	2	7.35	11	2	9.71	5	1	7.66	10	12	6.51	11	2	5.82	3
5	8.33	10	5	7.73	11	12	9.85	5	3	7.71	10	5	7.13	11	5	5.83	3
6	8.52	2	7	7.75	11	10	10.39	5	12	7.77	10	2	7.14	11	6	6.12	3
9	8.73	2	1	8.12	11	9	10.40	5	6	7.95	10	1	7.45	3	4	6.44	7
3	9.21	2 <sup>*4</sup>	8	8.18	11	3	10.67	5	5	8.40	10	8	7.50	3	3	7.67	7
10	10.08	-	3	8.71	11	⑥	10.69	5	⑨	8.76	-	11	7.50	3	11	8.06	7
⑧ <sup>*1</sup>	11.03	-	⑩	9.11	11	1	11.20	-	2	8.78	-	7	8.89	-	⑧	8.27	-
①	11.20	-	⑪	10.30	-	⑤	12.33	-	⑩	9.95	-	③	9.75	-	⑦	12.17	-
ONE WAY ANOVA																	
F (11,309) = 15.02			8.24			7.17			13.47			10.05			25.20		
P < 0.001			0.001			0.001			0.001			0.001			0.001		

\*1 CIRCLED SITUATIONS DEPICT THIS SOCIAL POWER TYPE

\*2 MEANS BASED ON THE SUM OF THREE BEST ITEMS FOR THIS SCALE (High # = High Power) (range: 3-15)

\*3 READ AS: Situation #5 is first situation significantly different from #7 ( $\alpha = .05$ )

\*4 NO SIGNIFICANT DIFFERENCE IN THE REMAINING SITUATIONS

comparison of the SP scales with the social influence components of the extended Fishbein model (Fishbein and Ajzen, 1975). This model is based on Dulany's (1961) verbal conditioning research and suggests that the amount of social influence is dependent upon the influencee's motivation to comply with the influencer. For example, one possible hypothesis that might be tested is that as the degree of referent social power increases, so will the motivation to comply. This analysis might provide insights into the reasons why one referent is more influential than another. In understanding family decision-making, this information might be particularly useful.

In conclusion, social power and influence are important aspects of many consumer decisions. To aid in the understanding of these aspects, this paper has attempted to provide a first step in the development of reliable and valid measures for the bases of social power.

#### References

- 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 (February, 1976), 3-11.
- Donn Byrne, "Interpersonal Attraction and Attitude Similarity," Journal of Abnormal and Social Psychology, 62 (May, 1961), 713-715.
- \_\_\_\_\_ and Don Nelson, "Attraction as a Linear Similarity," Journal of Personality and Social Psychology, (June, 1965), 659-663.
- Darwin Cartwright, "A Field Theoretical Conception of Power," in D. Cartwright (ed.) Studies in Social Power (Ann Arbor, Mich.: Institute for Social Research, 1959).
- \_\_\_\_\_ and Alvin Zander, Group Dynamics (New York: Harper and Row, 1968), 266.
- Barry E. Collins and Bertram Raven, "Group Structure: Attraction, Coalitions, Communication and Power," in G. Lindsey and E. Aronson (eds.) The Handbook of Social Psychology, Vol. 4 (Reading, Mass.: Addison-Wesley, 1969).
- Lee J. Cronbach, "Coefficient Alpha and the Internal Structure of Tests," Psychometrika, 16 (1951).
- Morton Deutsch and Harold Gerald, "A Study of Normative and Informational Influence Upon Individual Judgment," Journal of Abnormal and Social Psychology, 51 (1955), 629-636.
- Don E. Dulany, Jr., "Hypotheses and Habits in Verbal Operant Conditioning," Journal of Abnormal and Social Psychology, 63 (September, 1961), 251-263.
- Martin Fishbein and Icek Ajzen, Belief, Attitude, Intention and Behavior: An Introduction to Theory and Action (Reading, Mass.: Addison-Wesley, 1975).
- John R. P. French and Bertram Raven, "The Basis of Social Power," in D. Cartwright (ed.) Studies in Social Power, (Ann Arbor, Mich.: Institute for Social Research, 1959).
- Bert F. Green, "Attitude Measurement," in Handbook of Social Psychology, Vol. 1 (Reading, Mass.: Addison-Wesley, 1954).
- Harold H. Kelly, "Two Functions of Reference Groups," in G. E. Sevanson, T. M. Newcomb and E. L. Hartley (eds.) Readings in Social Psychology (New York: Holt, Rinehart and Winston, Inc., 1952).
- Helena Leet-Pellegrini and Jeffrey Z. Rubin, "The Effects of Six Bases of Power Upon Compliance, Identification, and Internalization," Bulletin of the Psychonomic Society, 3 (January, 1974), 63-70.
- Kurt Lewin, Field Theory in Social Science (New York: Harper, 1951)
- William McKelvey, "An Approach for Developing Shorter and Better Measuring Instruments," Working Paper #76-6, Human Systems Development Study Center, UCLA, 1976.
- Jun C. Nunnally, Psychometric Theory (New York: McGraw Hill Co., 1967).
- William E. Pollard and Terence R. Mitchell, "Decision Theory Analysis of Social Power," Psychological Bulletin 78 (November, 1972), 433-446.
- Bertram H. Raven, "Social Influence and Power," in I. D. Steiner and M. Fishbein (eds.) Current Studies in Social Psychology (New York: Holt, Rinehart and Winston, Inc., 1965), 371-382.
- \_\_\_\_\_, "A Cartoon Task for the Analysis of the Bases of Social Power," Unpublished manuscript, UCLA, 1973.
- \_\_\_\_\_, Richard Centers, and Aroldo Rodrigues, "The Bases of Conjugal Power," in R. E. Cromwell and D. H. Olson (eds.) Power in Families (New York: Sage Publications, 1975).
- Philip J. Runkel and Joseph E. McGrath, Research on Human Behavior: A Systematic Guide to Method (New York: Holt, Rinehart and Winston, Inc., 1972).
- Constantina Safelios-Rothschild, "The Study of Family Power Structure: A Review 1960-1969," Journal of Marriage and the Family, 32 (November, 1970), 539-552.
- John Schopler and Nicholas Bateson, "The Power of Dependence," Journal of Personality and Social Psychology, 2 (1965), 247-254.

## DEVIANT CONSUMER BEHAVIOR: A DIFFERENT VIEW

Michael K. Mills, University of Pittsburgh  
Thomas V. Bonoma, University of Pittsburgh

### Abstract

Previous research in the area of deviant consumer behavior in retail contexts has been theoretically and methodologically exploratory, with correspondingly little ability to get at the real problem--the motivation behind such behavior. A new theoretical perspective is offered which utilizes three prototypical power systems in a conflict based approach to the problem. Pilot tests of the theory's applicability to explaining theft problems are presented which show encouraging support for the theory.

### Introduction

Deviant consumer behavior is defined here as behavior in a retail store that society considers inappropriate or in conflict with a previously accepted societal norm. Some examples are shoplifting, price altering, destroying or damaging merchandise, marring in-store fixtures or restrooms, and consumer fraud. Such behavior is rampant and is on the increase. Indeed, one observer argues that the Gross National Dishonesty Quotient is rising steadily (Moneysworth, July 24, 1972).

Deviant consumer behavior involves a substantial cost to merchants. Stolen merchandise alone costs retailers between sixty and one hundred million dollars annually (Wright, 1972). F.B.I. figures show that the cost of recovered merchandise from arrested shoplifters was approximately 22 million dollars in 1975 (Uniform Crime Reports, 1975); the bulk, of course, goes unmeasured. The total estimated cost to retailers in shoplifting-related expenses totaled three billion dollars in 1959 (U.S. News and World Report, March 2, 1970). A conservative current estimate would place this close to twenty billion dollars annually.

In these inflationary times, stock "shortages" are a major marketing problem. Yet, there has been relatively little research done on deviant consumer behavior or the motivation behind such behavior (Cameron, 1964; Cohen and Stark, 1974; Geurts et. al, 1975; Pedrini, 1972). Additionally, these studies and other work in the area suffer from the lack of an actionable theoretical base.

### Theoretical Problems

Essentially all previous studies have approached the problem through labeling theory (Becker, 1960; Erickson, 1962), or have focused attention on the social psychological variables suggested by the deterrence literature (Andenaes, 1966; Gibbs, 1966). However, vast disagreement exists even among practitioners of each of these theoretical perspectives as to their applicability, or usefulness in explaining, controlling, or predicting deviant phenomena (Meade, 1974; Tittle, 1975). Further, neither labeling nor deterrence theory enjoys unequivocal or even firm support in the literature and, as some suggest, much of that which is cited in regard to the issue is of dubious value (Tittle and Logan, 1973).

Thus, we suggest that the theoretical foundation upon which many current views of deviant consumer behavior

are based is weak. Further, the competing theoretical orientations available allow for no integrated approach to the problem. Because of the increasing urgency of deviant consumer behavior to retailers and to marketers generally, a new theoretical perspective seems called for.

### A New Approach

This study utilizes a new theoretical perspective to explain deviant consumer behavior in retail stores, as well as the motivations behind such behavior. Basically it is our belief, supported by previous retail image research (Mills, 1976), that retail stores transmit social communications to their customers through their physical and marketing image. We believe consumers in "interaction" with stores react to such signals and modify their behavior accordingly. Thus, the deviant behaviors consumers engage in vis-a-vis retail stores are also social communications-generated in response to messages contained in store "image" manipulation.

This approach to the problem extends both retail image studies and a power-context interaction theory of Bonoma (Bonoma, 1976). In this view, consumers' perceptions of stores' relative power play an important role in understanding the motivations underlying both nondeviant and deviant consumer behavior. Research on such power perceptions may then suggest how power can be manipulated in social interaction (Bacharach and Lawler, 1976). Shoplifting, vandalism and other acts may thereby be removed from a nonactionable category of personal "character disorders" and put into the category of social communications, making them available for managerial control through increased attention to store image.

This general approach, novel in its application, lies on a solid theoretical and practical base. The concepts of conflict, cooperation, trust, power and social influence form a theoretical foundation that has generated much interest and has enjoyed far-ranging application in the social disciplines (Boulding, 1962; Deutsch 1973; Gamson, 1968). Bonoma (1976), however, has recently extended much of the power literature in arguing for a context-specific definition of power. He contends that a somewhat finer-grained examination of social episodes in which the constructs of power, conflict, cooperation and trust may be implicated shows clearly that these episodes fall not into one grand conceptual category, but into three. These are the unilateral, bilateral, and group welfare systems.

### Retail Power Contexts

We claim that the three systems outlined by Bonoma have their equivalents in retail store settings with corresponding implications for the analysis of deviant consumer behavior (Mills, 1976). For example, a large department store chain might make it explicitly clear to the consumer that the chain is strong and the consumer is weak by its size, by its strict adherence to a no-sale policy, by its tight credit policies, or by its appeal to one very small (e.g., upper-middle class)

segment of the retailing market. Another chain, contrarily, might encourage more equal comparisons between the consumer and itself by vigorous branching, by appeal to a wide segment of the target market, and numerous sales promotions. Even the provision or lack of parking may affect the power image transmitted to the customer. It is our claim that when stores are broken up in this fashion and put into one of the three theoretical categories, clear and consistent differences will be found in deviant consumer behavior occurring in various types of stores. The point is that store image generates impressions of store power in consumers; they may feel weaker, about equal, or a "part of" the selling organization, and will steal and vandalize more or less accordingly.

Evidence from previous empirical work would tend to support our contention that the power context of a store influences the degree of customers' deviant acts. In several studies (Kraut, 1976; Robin, 1963), highly ranked reasons given for shoplifting included the fact that the person wanted to get back at the store, had political motives, or felt shoplifting was acceptable.

Power is implicated further by a second factor which shows that most shoplifting and other deviant consumer behavior is done by juveniles (Cohen and Stark, 1974; Kraut, 1976; Robin, 1963) -- the "weaklings" of adult society. Further, rebellion, alienation, and powerlessness are important reasons given by apprehended juveniles for their behavior (Pedrini, 1972). Involvement of the young age group is demonstrated by the fact that 35 percent of larceny-theft crimes cleared in the nation's cities are solved by arrest of persons under 18 years of age (Uniform Crime Reports, 1975).

#### Hypotheses

The relationships specified above, the findings of previous research in the area and the researchers' best conjectures suggest the following theoretical assumptions and consequent hypotheses:

H<sub>1</sub> : The power context of a retail store as perceived by the customer and management can be identified and classified into one of our three types. Each type will differentially influence the incidence of deviant consumer behavior occurring in that store. Theoretical Justification: Retail stores can be classified as having more or less relative expendable influence resources vis-a-vis single consumers. Grossly speaking, those having the most power toward the unilateral ideal-point; those with the least, toward the group welfare.

H<sub>2</sub> : Perceived power differences between stores will vary as a consequence of various store attributes. Specifically,

H<sub>2a</sub> : Larger stores will be perceived as having a greater amount of power vis-a-vis the customers than smaller stores. Theoretical Justification: Size is an obvious indicator of resource base.

H<sub>2b</sub> : Stores with a greater number of branch operations will be perceived as having more power vis-a-vis the consumers than will stores having a fewer number of branches. Theoretical Justification: Same as for H<sub>2a</sub>, with the possible exception that vigorous branching may be perceived as in customers' best interests and hence, give impressions of

power equalization.

H<sub>2c</sub> : Stores catering to the elite or upper middle class shopper will be perceived as having greater power vis-a-vis the customer than will stores catering to the lower or lower middle class. Theoretical Justification: Stores capable of dealing with "powerful" consumers will be seen as more powerful than those dealing with "weak" consumers.

H<sub>2d</sub> : Stores with greater sales volume will be perceived as more powerful vis-a-vis the customer than will stores of lesser volume. Theoretical Justification: Dollar volume is an obvious indicator of resource base.

H<sub>2e</sub> : The greater the reliance on promotional campaigns or advertising by stores, the more will be the perceived power of the stores vis-a-vis the customer. Theoretical Justification: Promotional expenditures, like size and sales, indicate many disposable resources.

H<sub>3</sub> : The greater the customer's perceptions of power differences between store and customer, the greater will be the incidence of deviant consumer behavior as a counterpower move by consumers. Specifically,

H<sub>3a</sub> : The greatest perceived power differences and the greatest incidence of deviant consumer behavior should occur in the unilateral retail power system. Theoretical Justification: Implied by the theory and assumptions.

H<sub>3b</sub> : The least amount of perceived power differences and hence, the least incidence of deviant consumer behavior will occur in the bilateral retail power system. Theoretical Justification: Implied by theory and assumptions.

H<sub>3c</sub> : The amount of perceived power and the incidence of deviant consumer behavior of the mixed (bargaining) power system will fall somewhere in-between the unilateral and bilateral power systems. Theoretical Justification: Implied by the theory and assumptions.

H<sub>4</sub> : The retail power image of a given store will be manipulable by changes in the above attributes. Thus,

H<sub>4a</sub> : Rates of deviant consumer behavior will be manipulable by conscious setting of store image. Theoretical Justification: Necessary consequent of theory validation.

#### An Archival Pilot Test of the Hypotheses

As a partial test of the soundness of the hypotheses an archival-based preliminary pilot test was undertaken. Archival data consisting of stock shortage figures for retail department stores, gathered and tabulated by the National Retail Merchants Association and published in their annual volume, Financial and Operating Results, was examined and analyzed for the fifteen-year period 1961-1975. Stock shortages here serve as a proxy variable for shoplifting, since NRMA officials estimate

that over eighty percent of stock shortages represent theft. Several analyses were made, with the following results:

A chi-square analysis of levels of shoplifting to department store size was made; the six groupings listed in the raw data were combined into three groupings for analysis, and the criteria used to assign cell frequencies was based on low-to-high percentage values of shoplifting to sales for each year (this was done to assure adequate representation in each cell and to meet intuitive criteria). The null hypothesis that store size and level of shoplifting are independent was rejected at the  $p < .005$  level. Store size and level of shoplifting are related, with larger stores having a reliably greater incidence of shoplifting occurrences.

A chi-square was also done for the six original groupings of department store sizes represented in the Financial and Operating Results to level of shoplifting. The null hypothesis that store size and level of shoplifting are independent was rejected at the .005 level. Again, larger stores appear to have a greater incidence of shoplifting.

While only a first step, this pilot test revealed very significant and encouraging results as to the soundness of one of the hypothesized relationships--that of the relationship of store size (one contributing aspect of perceived store power) to levels of shoplifting.

#### Pilot Test 2: A Scenario Study

In order to get a "second opinion" on the validity of the theory which was not limited by the vagaries of an archival data base, we also conducted an experimental study at the University of Pittsburgh with college-student subjects.

#### The Instrument

We constructed six separate scenarios designed to operationalize one or another facet of the unilateral, bargaining, or bilateral power system within retailing establishments. The six scenarios were constructed in such a fashion that certain additional data, deemed useful commentary on the theory, could be collected besides a simple overall judgment of whether the theory could be generally confirmed by the test. Basically, the scenarios were as follows:

- Scenario A described a large "prestige" department store catering to customers in the upper-middle income ranges, whose pricing structure was moderately high, which advertised frequently, had its own credit cards, a downtown location, and somewhat difficult access. Scenario A typified a department store in the unilateral power context.
- Scenario B depicted a department store as large as Store A, but located in a suburban environment, catering to wide income segments, offering significant sales through promotional media, giving free parking, and so on. Store B was deliberately made to be a mixed description not characteristic of any of our three pure power-contexts.
- Scenario C was an automobile dealership, whose primary characteristic was the negotiable nature of the pricing structure. It was representative of a bargaining power-context in a nondepartment store setting.
- Store D represented a retail grocery cooperative in a university district. The store is owned by

its members, who also provide the labor to run the store. The store aims at the lower and lower middle income classes, is not-for-profit in structure, and uses only word-of-mouth advertising. Store D was intended to characterize a group welfare, power-context store in a nondepartment setting.

- Store E was a medium sized department store with a few branches, aiming at a full range of the income market, which sometimes negotiates for merchandise, which has liberal credit policies, which provides free parking, which does not advertise frequently, and depends on word-of-mouth about its good bargains to bring in customers. Store E was designed to illustrate a bargaining power-context within a department store setting.
- Finally, store F was characterized as a small recreational clothing cooperative. Store F was designed to characterize a bilateral power-context department store setting.

Choice of these six scenarios gave us a number of interesting comparisons to perform including

- (1) An overall comparison between the unilateral power-context scenario (Scenario A), and the bargaining (Scenarios C and E) and bilateral (Scenarios D and F) contexts,
- (2) A within department store only comparison,
- (3) An across retailing establishment comparison using say, a department store (Scenario A) for the unilateral power system, an auto dealer (Scenario C) for the bargaining system, and a vacationer's buying cooperative (Scenario D) for the bilateral system, and
- (4) What we termed a "Big Chicken" test of a possible alternative account for our archival results. A colleague suggested that store size was possibly the only mediator of higher or lower levels of deviant consumer behavior. That is, he suggested the hypothesis that the bigger you are, the more likely you are to be the target of deviant consumer behavior. We not very kindly named this notion the "Big Chicken" theory, and set about to test it in this pilot study by constructing a scenario B which displayed a large department store with characteristics of each of the unilateral, the bargaining, and the bilateral systems, so that we could check his speculations against the pure types of systems depicted by the other scenarios.

#### Subjects and Method

The subjects were 66 MBA and undergraduate business students at the University of Pittsburgh. The sample was split about equally between graduate and undergraduate students, and about equally between male and female subjects. Subjects volunteered for the study with no course credit and no other incentive for completion of the survey.

Subjects were approached on an in-class basis, and asked to take a copy of the "Retail Store Survey" home with them during the month of January, 1978. They were asked to complete the survey and to return it to a central University location by 1 February, 1978. Eighty surveys were handed out in this manner; 66 were returned as promised. Each subject received a booklet in which the six scenarios described above were placed in randomized order. This was done to minimize the effects of sequence on their responses.

Dependent Measures

Each scenario was followed by a list of fifteen questions, broken up into three types. The manipulation check section asked five questions about the type of store depicted in the scenario, or about S's relationship between the store and himself.

The remaining questions asked subjects whether they had ever destroyed or damaged merchandise in this type of store, whether merchandise was ever returned fraudulently to the store, whether shoplifting occurred in this kind of store, whether this kind of store would be vandalized, and whether fraudulent complaint would be made in this type of store. The questions were asked in two different ways. One set of questions asked whether the subject ever engaged in these actions. The other set of questions asked how likely subjects thought it would be that "someone" would engage in the behaviors so described. The "self" questions contained only a "yes," "no," or "doesn't apply" response listing. The second, "someone" set of questions, contained a five point Likert scale from very unlikely to very likely.

We included these "self" and "other" questions because we regarded it as highly unlikely, without a randomized research design, that subjects would actually report they had shoplifted, vandalized, or engaged in other fraudulent behaviors in a department store. Yet, it was our belief that subjects would be likely to endorse that "someone" might engage in such behaviors, and we were willing to accept this data as a preliminary commentary on the theory.

The "self" questions were combined to form a summary index which could range from 0 (all "no" or it "doesn't apply") to 5 (all "yes"). The "someone" questions were also summed to form a composite index varying from 0 (all "very unlikely") to 5 (all "very likely").

The three summary indices - one for the manipulation check, and one each for the "someone" and "self" questions - were subjected to a one-way repeated measures analysis of variance. Planned comparisons, consistent with the description given above, were performed on means generated for each of the scenarios.

Results

Tables 1 and 2 respectively, present the results for the manipulation check and the "someone" results. The "self" results, as expected, yielded generally insignificant results (though in the direction predicted by the theory), and to save space are not reported here. Turning to Table 1, an assessment can be made of whether our somewhat casually constructed scenarios did indeed operationalize the three power contexts of unilateral, bargaining, and bilateral reactions. The F-test computed on the summary of the five manipulation check questions detailed in Table 1 yielded a highly significant F-ratio of 40.16, reliable at the .0001 level.

The first check that was made on the individual scenario means comprising this F-ratio was an overall comparison. The results indicated that we were indeed successful in operationalizing the unilateral, bargaining, and bilateral systems. As Table 1 shows, the unilateral scenario generated the highest value on our summary index, the bargaining and intermediate value, and the bilateral system the lowest value. All differences (by two-tailed t-tests) were significant at the .05 level.

When the manipulation check was completed within department stores only, and across retailing contexts,

TABLE 1

PILOT 2: RESULTS ON MANIPULATION

- Manipulation Check: "Does/Is this store (1) make too much money (agree = 5 - disagree = 1); (2) very powerful relative to yourself (5-1); (3) generate a lot of conflict between it and you (5-1); (4) allow you to influence stocked items (1-5); (5) allow you to correct wrong treatment to yourself (1-5)?"
- Summary index from 5-25 computed.
- F(1, 65) = 40.16 (p < .0001).

PLANNED COMPARISONS

- (1) Overall:
 

Unilateral = Scenario A = 15.3	.05	}	.05
Bargaining = C & E = 13.5	.05		
Bilateral = D & F = 11.5			
- (2) Within Department Stores Only:
 

Unilateral = Scenario A = 15.3	ns	}	.05
Bargaining = Scenario E = 14.4	ns		
Bilateral = Scenario F = 12.0			
- (3) Across Retailers:
 

Unilateral = Scenario A = 15.3	ns	}	.05
Bargaining = Scenario C = 15.0			
Bilateral = Scenario D = 11.0	.05		
- (4) "Big Chicken" Theory:
 

Unilateral = Scenario A = 15.3	}	ns
Mixed Big = Scenario B = 14.4		

- The Manipulation Check Was Successful: Ss perceived "unilateral" store more powerful, conflictful, and making too much money. Ss perceived themselves as having less influence and less ability to redress wrongs than in either bargaining or bilateral stores.

TABLE 2

PILOT 2: "SOMEONE" RESULTS

- How likely in this type of store that someone would
  - destroy merchandise (1=not likely-5=very likely)
  - fraudulently return merchandise (1-5)
  - shoplift (1-5)
  - vandalize (1-5)
  - fraudulently complain (1-5)
- Summary index from 5-25 computed
- F(1, 61) = 29.09, p < .0001

TABLE 2 (Contd.)

PLANNED COMPARISONS

(1) Overall:

Unilateral = Scenario A	15.6	] ns	} .05
Bargaining = C & E =	15.3	]	
Bilateral = D & F =	12.9	] .05	

(2) Within Department Stores Only:

Unilateral = Scenario A =	15.6	] ns	} .05
Bargaining = Scenario E =	17.6	]	
Bilateral = Scenario F =	13.2	] .05	

(3) Across Retailers:

Unilateral = Scenario A =	15.6	] .05	} .05
Bargaining = Scenario C =	13.6	]	
Bilateral = Scenario D =	12.6	] ns	

(4) "Big Chicken" Theory:

Unilateral = Scenario A =	15.6	} .05
Mixed Big = Scenario B =	17.9	

- The results for "someone" generally confirm the theory. Subjects are more likely to report someone will shoplift, vandalize, etc. unilateral than bargaining than bilateral stores.

the results again suggested that we for the most part successfully operationalized the three power systems. Even though some of the individual planned comparisons were nonsignificant, all means were in the predicted direction.

Finally, to check whether an equally large but not unilaterally-disposed department store generated higher or lower (our prediction) perceived power than a strictly unilateral department store, we compared the mean for scenario A against that with scenario B. As Table 1 shows, the perceived power score for the mixed type department store was indeed lower than that for the strictly unilateral department store. However, the results were non-significant.

The results on whether subjects would endorse that "someone" would destroy merchandise, fraudulently return it, shoplift, or vandalize, or fraudulently complain in one or another of the store scenarios were considered to be the main ones of this study. Again, an overall comparison, a within department stores only, an across retailers comparison, and a "Big Chicken" comparison were made on the scenario means.

Over all the scenarios broken down into unilateral, bargaining, and bilateral power types, the means were in the theoretically predicted direction. Subjects reported that "someone" would more likely destroy merchandise, fraudulently return it, shoplift, vandalize, or fraudulently complain in the unilateral power system. They endorsed such behavior less extremely in the bargaining system, and least in the bilateral system. However, all means were not significant in planned comparisons. Only the unilateral/bilateral and bargaining/bilateral comparisons were significant. Nonetheless, the results must be taken as encouraging for the theory.

Within department stores only, a similar set of results was obtained. That is, the only reliable planned comparisons were between the unilateral and bilateral systems, and the bargaining and bilateral systems. In each case, the bilateral system elicited a lower rate of endorsement for these deviant consumer acts than did the other system. The unilateral-bargaining comparison was nonsignificant, although the bargaining scenario did generate a slightly higher level of endorsement of deviant consumer acts, a result which is contrary to the power-context theory.

Across retailers, the means were again all in the theoretically predicted order of unilateral > bargaining > bilateral. Again, only two of the three possible comparisons were significant. This time, the bargaining-bilateral comparison was not reliable, although the other two comparisons were.

Finally, the test of the "Big Chicken" theory, in which a strictly unilateral power-context store was pitted against one equally large but mixed in its characteristics, yielded a theoretically disconfirming result. That is, both stores were equally large, but the one in a suburban location, with free parking, with a more "bargaining oriented" base generated higher levels of endorsement for the deviant consumer act cited than the strict unilateral power-context department store. This result lends some credence to the argument that size alone may be an important factor in whether or not subjects would behave deviantly within retail contexts.

Discussion

The results of our second pilot study, conducted with business students responding to scenarios operationalizing the three power systems, lend some strong support to the power-context theory. When combined with the results of the first pilot test, a strong case can be made that the power-context theory appears to offer a rational explanation for deviant consumer behavior, and is worthy of more controlled and complete tests under more realistic circumstances.

Implications of the Approach

While the conceptual approach we have developed has yet to be more than tested in an exploratory way, it has potentially far-reaching implications on theoretical and practical grounds.

Theoretical Implications

Given support for our hypothesized relationships, much theoretical integration occurs. Support for the hypothesized relationships would offer a viable alternative to the labeling versus deterrence chaos that now exists in studying consumer deviance. The point is not that existing theories of deviance must be abandoned. Rather, they may need to be modified to include principles of "context translation" which explain how the theoretical constructs may be modified across various power systems.

Practical Implications

There are also practical benefits to be gained here, with a number of parties who stand to benefit from further research. One such party is the nation's retailers. If the power difference in a store-customer relationship affects deviant consumer behavior, this suggests enormous challenges to retailers in the development of a favorable retail image. Among other things, it would certainly suggest a "lowering of posture" or more efforts at conscious image building, or similar

attempts to play down perceived store-customer power differences. Such modifications should result in decreased incidence of deviant consumer behavior resulting in increased profits for retailers.

Though precise specification of managerial image-control implications will have to wait for future research, we can pose some illustrative research questions: What are the effects of an "information hotline" in the store in reducing perceived power disparities, and hence, deviant behavior? Should a single branch-store be put in the downtown area, where it is accessible (good) but parking is impossible (bad), or should it go in the relatively inaccessible but less dense suburbs? What are the effects of a "cash-only" or "house credit only" policy on shoplifting? Do armed guards encourage shoplifting, or do they retard it?

#### Conclusion

If our view of deviant consumer behavior is supported by more data, then several advances may be toted up for the area in general:

- An integrated, interactive theory of deviant consumer behavior will be available for further development and refinement.
- The "onus" of deviant consumer acts like shoplifting will be shifted from the consumer's narrow shoulders to the joint interaction of the consumer and the retailer.
- A set of actionable and directly controllable variables, namely the marketing mix, will be identified as the major mechanism by which deviant consumer behavior can be reduced by the retailer.
- A great deal of money and other resources may be able to be partially reallocated from security and theft-control mechanisms toward more effectively meeting consumer needs at a more competitive price.

#### References

- J. Andenaes, The General Preventative Effects of Punishment, University of Pennsylvania Law Review, 114, (1966) 949-983.
- S.B. Bacharach and E.J. Lawler, "The Perception of Power," Social Forces, 55, (1976), 123-134.
- H. Becker, Outsiders: Studies in the Sociology of Deviance, (New York: Free Press, 1960).
- T.V. Bonoma, "Conflict, Cooperation and Trust in Three Power Systems," Behavioral Science, 21, (1976), 499-514.
- K.E. Boulding, Toward a Theory of Peace. In R. Fisher (Ed.), International Conflict and Behavioral Science, (New York: Basic, 1964).
- M.O. Cameron, The Booster and the Snitch, (New York: Free Press, 1964).
- L. Cohen and R. Stark, "Labeling Theory and the Five-Finger Discount," Journal of Criminal Justice, 2, (1974) 23-36.
- M. Deutsch, The Resolution of Conflict. (New Haven: Yale, 1973).
- K. Erickson, "Notes on the Sociology of Deviance," Social Problems, 9, (1962), 307-314.
- W.A. Gamson, The Strategy of Social Protest. (Homewood, IL: Dorsey, 1975).
- M.D. Geurts, J. Reinmuth, and P.R. Andrus, "Researching Shoplifting and Other Deviant Consumer Behavior, Using the Randomized Response Research Design," Journal of Retailing, 50, (1975), 43-48.
- J.P. Gibbs, "Sanctions," Social Problems, 15, (1966), 14, 147-159.
- R.E. Kraut, "Deterrent and Definitional Influences on Shoplifting," Social Problems, 23, (1976), 358-368.
- D. Matza, Becoming Delinquent. (Englewood Cliffs: Prentice Hall, 1969).
- A.C. Meade, "The Labeling Approach to Delinquency: State of the Theory as a Function of Method," Social Forces, 53, (1974), 83-91.
- M.K. Mills, "The Effects of Perceived Conflict, Power, and Influence on Retail Store Image: An Exploratory Study in Three Power Systems," unpublished manuscript, Graduate School of Business, University of Pittsburgh, 1976.
- Moneysworth, July 24, 1973, 3.
- T. Parsons, On the Concept of Influence. In T. Parsons (Ed.), Sociological Theory and Modern Society, (New York: Free Press, 1967).
- D.T. Pedrini, "Psychology and the Problems of Shoplifting." St. Louis, MO: Paper presented at the Midwest Section of the American Business Association Conference, April, 1972.
- "Rising Wave of Shoplifting," U.S. News and World Report, (March 2, 1970), 56-58.
- G. Robin, "Patterns of Department Store Shoplifting," Crime and Delinquency, 9, (1963), 163-173.
- C.R. Tittle, "Deterrents or Labeling?," Social Forces, 53, (1975), 399-410.
- C.R. Tittle, and C.H. Logan, "Sanctions and Deviance: Evidence and Remaining Questions," Law and Society Review, 7, (1973), 372-392.
- Uniform Crime Reports (Washington: Federal Bureau of Investigation, 1975).
- R.A. Wright, "Nation's Retail Merchants Mobilize Security Systems to Combat Fast-Growing Shoplifting Trend," New York Times, 121, (1), 51.
- N.B. Zabriskie, "Fraud by Consumers," Journal of Retailing, 48, (1972), 22-27.



A COMMENT ON VALUE STRUCTURES, POWER-CONTENT INTERACTION THEORY,  
AND THE MEASUREMENT OF SOCIAL POWER

Stewart W. Bither, The Pennsylvania State University

Abstract

The purpose of this paper is to review three studies presented under the session title Psychological Contributions. In fact the papers range from a psychological through a social psychological to a sociological perspective. As one might suspect from these perspectives, the papers are really quite unrelated. Each does present an interesting and perhaps provocative idea. Each also represents exploratory type research which is in keeping with conference goals of encouraging discussion of new research directions. Due to the lack of common theme among the papers it seems appropriate to review each separately and comment about future directions in terms of the distinct area each paper represents.

Value Structures

The paper "Value Structures and Consumer Behavior" by Gutman and Vinson begins with the statement that personal value systems play an important role in consumer behavior. The central proposition of the paper is that the explanatory power of the value construct has been limited by over-simplified conceptualizations of value structure and by cumbersome instruments for measuring the structure. Specifically Gutman and Vinson state that we must go beyond the basic research conceptualization of an interconnected structure of values with a central-peripheral dimensionality. They suggest "a more adequate system would be one in which some values are seen as more consistent with one another to the effect that the same behaviors are instrumental in achieving them." They further quote Williams (1968) in noting that "...particular acts or sequences of acts are steered by multiple and changing clusters of values."

This appears to be a viable and potentially productive way of conceptualizing values in that different behavior may be seen as relating to different clusters of value. Thus, depending upon the situation any two values may be viewed as consistent in that the behavior will likely be instrumental in achieving them both, contrary in that achieving one will lead to blocking achievement of the other, or they may be viewed as totally independent in the context. Unfortunately, the methodology utilized in the research does not allow us to examine full scope of the relationships posited.

In the study a generalized perceptual map was constructed using dissimilarity measures taken from pairs of values from Rokeach's terminal values scales (Rokeach, 1968). In addition, instrumentality ratings of ten different behaviors for achieving or blocking each of the values were obtained. The behavior measures were then related to the generalized value map using direction cosines. The authors indicate that the INDSCAL solution in three dimensions results in a disappointing 36% variance accounted for and stress values are not given. More importantly, the basic perspectives that values may be viewed as consistent, opposed, or independent in terms of behavior is not tested by this methodology. In order to make such a test dissimilarity scales could be administered in the context of specific behaviors in order to compare the resultant preference spaces. Alternatively, a variety of experimental designs can be envisaged in which different behaviors could be related to clusters of values using relatively simple scaling procedures.

Power-Content Interaction Theory

The paper "Deviant Consumer Behavior: A Different View" by Mills and Bonoma suggests a theoretical perspective for understanding consumer deviant behavior. Deviant behavior is described by the authors as socially unacceptable behavior such as destructive behavior in the retail institution and all types of consumer fraud. The theoretical perspective is based upon the authors' belief that consumers modify their behavior in response "...to messages contained in store 'image' manipulation." The authors suggest that stores can be classified into three conceptual categories which differ in terms of power, conflict, cooperation and trust implications for patrons. This conceptualization follows Bonoma's (1976) classification of social power into three context specific categories: unilateral, bilateral and group welfare. The authors speculate that deviant behavior will be greatest when the store is seen as having unilateral power and least as the power base moves beyond the bilateral stages to the group welfare state. They believe that stores will be perceived as being powerful to the extent they are larger, have more branches, have greater sales volume, advertise more and cater to upper class groups. A final hypothesis suggests that image may be modified and deviant behavior rates changed by manipulating store image.

The study consisted of two parts. The first, an "archival" study which showed that large stores had a greater percentage of stock shortages than did small stores. The second part comprised a study in which scenarios representing the three conceptual categories were presented to undergraduate and graduate students who were asked the degree to which they believed deviant behavior would take place in each.

The conceptual base for this study is fascinating and the work on alienation would support some of the hypothesis base. In addition, there is a certain amount of face validity to the hypothesis base. Who has not experienced the utter frustration of fighting the bureaucracy in a large retail institution when there has been a foul up in computer billing or when other problems have occurred and I dare say that fleeting thoughts of fraud or worse have occurred to us all.

The study is very much an exploratory one and has served the purpose of raising a few alternative explanations which must be eliminated in future studies. The authors undoubtedly would be first to admit that attributions about deviant behavior made by university graduate (and undergraduate students do not constitute a strong support base for their hypotheses. Even if this were not the case, the scenarios used in the study are problematic. Different kinds of retail institutions e.g., a large department store versus an automobile dealership, represent different levels of opportunity to engage in various kinds of consumer fraud. Thus, the judgment that individuals might engage in fraud due to differentials in perceived power must be somehow tempered by the relative opportunity to engage in fraudulent activities across institution. A second problem with the scenarios is that by describing location and facilities of the department stores it is likely that respondents made some judgment about who the customers might be and attributed the likelihood of fraud not to the power base of the store but to the perceived age or socio-economic status of the customer base of the store. To the extent possible, these sources of alternative explanations should be eliminated in further studies.

In the "archival study" inventory shortages are equated with shoplifting because studies are cited which indicate that 80% of inventory shortages are due to theft. While shoplifting may be a large part of the theft, certainly employee theft and other non-consumer theft must constitute some part of the shortages. In addition, this part of the study is open to many reasonable explanations which are alternative to the power base one. Large retail institutions are more difficult to control, large retail institutions are often located in large cities where alienation levels many differ from smaller cities and so on.

This study demonstrates many of the problems of applying sociological concepts but offers a substantive contribution in terms of the basic constructs developed by the authors. As an exploratory study, it makes steps in the direction of construct clarification and ultimately validation. I found it a pleasure to read and consider the overstatement in both theory and results section of the paper to be a productive challenge for others in the field to become more involved in studying the social context of consumer behavior.

#### The Measurement of Social Power

The paper "Measuring the Bases of Social Power" by Swazy reports an effort to build and validate a scale measuring French and Ravens (1959) six bases of social power. Swazy briefly reviews the power bases; reward, coercive, referent, legitimate expert and information power, and then describes in some detail the scale construction and validation procedures he has undertaken. The scaling procedure used was the Likert scale. A pool of 150 scale items representing various aspects of the power bases was generated by the author and judged by six judges as to its appropriateness to the 6 power bases. From this procedure 85 scale items were selected. Scenarios representing depictions of pure type social power situations were generated in much the same manner. Although it might have been preferable to generate item pools from a panel of judges to gain a more comprehensive pool of power items this is not a major problem.

In the study, 85 final scale items were rated by a sample of 321 undergraduate students. Each student rated all 85 scale items in terms of one of 12 scenarios. Items were selected for the final scale by choosing those with high discriminating power for the scenarios presented. This was accomplished by choosing scales with the highest total common variance with the uni-factor solutions using a factor analytic procedure recommended by McKelvy, 1976. Measures of internal consistency, Cronback's (1951) alpha, were high among the selected scales items.

A test of the discriminating ability of the scales was accomplished by calculating the scores across all scenarios for each of the six three item scales. The statistical procedure used was a one way ANOVA. Newman Kuels tests of mean differences were also presented. In general, the discriminating ability of the scales was good. There appear, however, to be major problems with the legitimate power scale.

This study seems to have been carefully done and the results clearly described. It constitutes an important first step in operationalizing measures for the six social power bases. The study is both exploratory and developmental in that the scale items were selected based upon an intercorrelation matrix derived from the data on which the scales were ultimately tested. The next step would be to test the scales against a different set of power scenarios generated in a manner similar to that described in the paper. This would give us more confidence in the discriminatory ability of the scales.

Additional work needs to be done on deriving an acceptable legitimate power scale. The data suggest that more than one construct may be included in the definition or operationalization of this scale. Concurrently, additional validation work could be undertaken by repeating the factor analytic procedures used in scale constructions on different scenarios in order to get a better measure of reliability and internal scale consistency.

This study appears to be a sound beginning to a more thorough examination of the constructs involved in explicating the social power concept. However, the construct validation process will be a difficult one. I expect we will have great difficulty in identifying "pure" examples of individual social power bases even in contrived laboratory situations. Thus, we may become involved in assessing the degree to which various social power bases are perceived as being present in an environment. If this turns out to be the case, and I consider it likely, then we will need a great deal more work in operationalizing the definitions of the social power bases. The scales developed in this paper discriminate among scenarios chosen to represent pure power base type situations. No attempt at all is made to measure the degree to which a power base is present or absent.

#### References

- T. V. Bonoma, "Conflict, Cooperation and Trust in Three Power Systems," Behavioral Science, 21, (1976), 499-514.
- Lee J. Cronbach, "Coefficient Alpha and the Internal Structure of Tests," Psychometrika, 16 (1951).
- John R. P. French and Bertram Raven, "The Basis of Social Power," in D. Cartwright (ed.) Studies in Social Power (Ann Arbor, Mich.: Institute for Social Research, 1969).
- William McKelvey, "An Approach for Developing Shorter and Better Measuring Instruments," Working Paper #76-6, Human Systems Development Study Center, UCLA, 1976.
- Milton Rikeach, "A Theory of Organization and Change Within Value-Attitude System," Journal of Social Sciences, 24 (1968), 13-33.
- R. M. Williams, International Encyclopedia of the Social Science (New York: Macmillan, 1968), p. 287.

## DISCUSSION OF PSYCHOLOGICAL CONTRIBUTIONS

Jagdish N. Sheth, University of Illinois

### Abstract

The three research papers on deviant behavior, structure of human values, and measurement of social power have one thing in common: They are all psychological contributions to the consumer behavior area. After evaluating each paper, several suggestions are made about the future directions in each area with respect to theory development, methodological issues, and testing propositions in the consumer behavior context.

### Introduction

At a first glance, the three papers seem to have very little in common except that each is a psychological contribution to the area of consumer behavior. Even this may be argued as only partially true because the same issues are equally researched in sociology and to some extent in consumption economics. However, a more careful look at the three research papers reveals at least four distinct common attributes.

First, all of them are anti-attitudinal papers and explicitly deviate from the more common multiattribute attitude research. This is remarkable in light of the extreme popularity of multiattribute models in consumer behavior, and may indicate that the tide of attitude research may be receding in consumer behavior.

Second, all of them are direct borrowings from psychological and sociological disciplines. However, unlike previous blind borrowings from the behavioral and social sciences, each paper has attempted either to critically examine what is theorized or measured in psychology or add richness of thinking and measurement to the original contributions. For example, Mills and Bonoma strongly dispute the labeling theory or deterrence research as explanations for the deviant consumer behavior, and they offer a new approach in terms of relative balance of power between the buyer and the seller in the market place. Similarly, Gutman and Vinson propose and test that Rokeach's Value Survey consists of many values which are not mutually exclusive but cluster together either in a complimentary or in an inverse manner. Finally, Swasy attempts to develop a psychometric scale with which to measure the six-factor power theory proposed by French and Raven.

Third, all the three papers deal with more fundamental and general constructs reminiscent of the golden days of personality and motivation research. Thus, Gutman and Vinson rely on fundamental human values as a relevant construct in consumer behavior and both Swasy and Mills and Bonoma rely on social power and balance of power as the relevant constructs in consumer behavior.

Fourth, all of them are pilot studies on a convenient sample of college students, and utilize a survey research approach. Of course, the sample sizes and specific instruments vary from one study to another.

We will first evaluate each research paper in terms of its strengths and weaknesses and then offer some suggestions for future research and theory in each area.

### Mills and Bonoma Study

In light of the recent zeal and almost missionary approach toward protecting the consumer against the malpractices of the market place, it is refreshing to note that someone has attempted to research and provide a theoretical foundation on shoplifting, vandalism and deception practiced by consumers against the marketer.

### Strengths

The biggest strength of the paper is the choice of the problem area. Deviant consumer behavior especially among the young adults has reached crisis proportions as a part of the overall syndrome of juvenile delinquency. Unfortunately, sociologists and psychologists have ignored the consumer behavior aspects in favor of deviant social behavior creating the need for us to research the area.

A second area of strength in the paper is the proposed context-specific balance of power approach as a way to understand and predict deviant consumer behavior. The classification of retail stores having images of unilateral, bilateral, and the group welfare power structures is a new approach to retail store image research.

A third strong point about the paper is its attempt to actually measure and test specific propositions derived from the theory in the retailing context. As I have commented elsewhere (Sheth, 1974), it is almost mandatory that authors of new theories also develop test instruments and hopefully carry out some empirical-experimental research on their theoretical perspectives.

### Weaknesses

Unfortunately, this research paper has more weaknesses than strengths. I will summarize the weaknesses in terms of three categories: conceptual, methodological, and analytical weaknesses.

Conceptually, it is very difficult to agree that all deviant consumer behavior is a function of the perceived balance of power between the store and the consumer. It is much more logical to suggest that deviant consumer behavior is a function of the interaction of three distinct factors: Opportunity to shoplift, vandalize or deceive the store, personal and socioeconomic characteristics of the consumer, and the perceived balance of power. Without such a complex three-way interaction model, it is not possible to explain why (a) some consumers never shoplift from any power base store, (b) some weak power stores such as the "mama and papa" stores get vandalized and held up, and (c) stores with comparable sizes, sales and power structure have different rates of shoplifting, vandalism and deception by consumers.

In addition, there are two other conceptual weaknesses. First, why should there be only three types of power systems? There is no rationale about the typology in the paper, and I presume that it is fully detailed in the Bonoma paper referred to in the study. Second, several subhypotheses are highly collinear and, therefore, unnecessary to be tested. For example, larger stores have greater sales volumes and spend more on promotional campaigns. Given this multicollinearity among size, sales and promotion, it is unnecessary to have three separate tests of hypotheses.

The methodological weaknesses center around the sampling and the data collection procedures. It is unfortunate that despite the national statistics about juveniles being the main segment of deviant consumers and its use to justify the power theory proposed by the authors, they chose to sample university students who are considerably older and more mature. Similarly, the authors chose to use the scenario method which is not context specific enough. There are too many other specific factors besides the general stereotypes which attract deviant behavior to a specific retail store. In fact, this is somewhat incongruent in light of the heavy emphasis the authors give to making the power theory more context specific.

The analytical weaknesses can be specified with respect to each of the two pilot studies. The first pilot study involved a chi-squared analysis of the relationship between size of the department store and level of shoplifting. Since shoplifting incidence was measured as a percentage of sales, and store size is more or less a collinear variable of sales, the rejection of the null hypothesis is more due to contamination of the dependent-independent variables than a true relationship between level of shoplifting and store size. It is also unfortunate that the amount of covariance contained between the two variables is not reported in the study.

In the second pilot study, the results of scenario manipulations are highly tentative. Since most of the mean values range between 11.0 and 15.0 on a scale of 5 to 25, it is obvious that the statistical significance is achieved by sample size rather than by amount of variance explained. In other words, rejection of the null hypothesis is not sufficient to accept that the experimental hypothesis is proved or successfully tested. We need to compare the results not against null hypotheses but against a priori hypotheses (Overall and Klett, 1974).

In summary, however, Mills and Bonoma must be applauded for pointing at a very relevant but neglected problem area of consumer behavior, and for making early efforts at both theorizing and testing the phenomenon of deviant consumer behavior.

#### Gutman and Vinson Study

This paper is an attempt to critically examine the structure of values as proposed by Rokeach. While the authors seem to accept the Rokeach theory of values and its relevance to consumer behavior, they are reluctant to accept the rank ordering of 36 terminal and instrumental values proposed in the Value Survey as a method of measuring values. They believe that many of the values are highly collinear and, therefore, cluster together and possess fewer underlying dimensions. In order to understand this clustering and reduced dimensionality of values in the Value Survey, they utilize the multidimensional scaling (MDS) approach.

#### Strengths

The primary strength of this paper is to point out weaknesses of the Rokeach's Value Survey instrument which had been utilized in the past without a careful methodological look and in a straight-forward manner in consumer behavior studies. The authors very correctly point out that the instrument is very difficult to administer and obtain error-free responses because of similarities of many values and the rank ordering procedure recommended by Rokeach.

It is indeed refreshing to note that researchers in consumer behavior have the confidence and maturity to question theories and measurement instruments proposed by

psychologists rather than to blindly borrow them as was so typical of the era of personality research in marketing.

#### Weaknesses

The choice of MDS as a method to understand clustering and dimensionality of values puzzles me. It is simply a very complicated procedure requiring numerous compromises in data collection and data analysis to warrant its usage for clustering purposes (Sheth, 1976). Most of my negative comments stem more from the use of MDS as a method to recover underlying dimensionality of Rokeach's values.

First, the problem of rank ordering of values is compounded in the MDS approach since it requires ranking of pairs of values. Thus, to obtain distances or similarities among 36 values in the Rokeach scale would require a total of 666 stimuli to be ranked which is admittedly a more difficult task than ranking 36 values. In fact, MDS users have often resorted to fractional factorial designs to reduce the comparison task. Even in this study, the authors limited themselves to only 14 of the 36 values in order to make the task of making judgments manageable. Furthermore, they were forced to use a distance rating scale for each pairwise combination rather than the ordinal rank ordering procedure because of still very large number (91) of pairs of values. The use of MDS, therefore, seems to be inconsistent with their rightful criticism of rank order requirement in the Rokeach scale.

Second, the three-dimensional MDS map cannot be generalized to the full set of values since it is derived from only a subset of 14 values. Its dimensionality is likely to increase as more values are added.

Third, it is not always possible to attribute the dimensions as due to perceived similarity of stimuli. The dimensions recovered in the MDS are often reflections of sample heterogeneity and dimensionality of a single stimulus. Both of these seem to be present in this study as indicated by large ranges of subject salencies for each of the three dimensions, and by mixed loadings of a value on two or more dimensions (Anderson and Sheth 1976).

Fourth, the three-dimensional INDSCAL solution is not at all meaningful despite the brave efforts by authors to find a structure in the data. For example, the first two dimensions do not have a single value loaded above 0.467 which is clearly a poor fit between values and the underlying dimensions. The third dimension has only one value loaded high enough to justify interpretation of that dimension in terms of clustering of values. As Redinger and Sheth (1977) have pointed out, this is to be expected when the data collection in MDS requires too many pairwise comparisons and the distance scale used is restricted to a seven point scale as was the case in this study.

Finally, except for the fourth behavior (getting high), all other behaviors have very comparable average rating as perceived instruments to attaining or blocking all of the 14 values. In other words, there is no real discrimination in terms of certain clusters of values being attained or blocked by certain behaviors as one should expect if values are useful in social or consumer behavior.

Rather than using MDS, the authors would have been considerably better off in deriving value structures if they had relied upon factor analysis or cluster analysis of the distance matrix between values. In fact, structure of values could be also revealed by simple or partial correlations among values in terms of their desira-

bility or saliency to the individual.

### Swasy Study

The Swasy paper on measurement of social power is an excellent example of psychometric scaling and scale development. This paper has far more strengths than weaknesses and, therefore, it is a pleasure to discuss it.

#### Strengths

The most important strength of this paper is the careful and rigorous procedures followed by Swasy in operationalizing, measuring and validating the Social Power scale. There is a very succinct and clear description of French and Raven typology of social power, an excellent critique of past efforts at operationalizing the construct, and a superb description of the procedures followed in deriving the social power scale. The author clearly has a very balanced background in both social psychology and testing and measurement which is often rare in psychology.

Swasy is also a very articulate writer. His paper reads with complete clarity and he is equally good in describing both theoretical and methodological aspects of his study.

#### Weaknesses

As I mentioned before, there are only minor weaknesses in this paper. Most of them relate to author's assertions about the goodness of fit between his theory and the data.

First, the amount of total variance retained by all factors seldom exceeds 53 percent. Similarly, the amount of variance summarized in the general factor is seldom above 35 percent. Those are simply very low percentages. It indicates that the item overlap is not as strong as one would expect in scale development. My suspicion is that this is due to heterogeneity of his sample despite the fact that they are all undergraduate accounting majors at UCLA (Sheth and Armstrong, 1969). Alternatively, it suggests that there is a lot of response error in the data among his subjects which is highly ad hoc and specific to each respondent.

A second weakness is in the validation stage. The results are not as clear-cut as Swasy would make us believe with respect to specific situations reflecting each type of social power base. For example, we should expect a true bimodal distribution of all situations on a particular social power base with mean values dichotomized at or near the extremes of the scale range between 3 and 15. The actual distributions of the mean values across all six social power bases are much narrower, and represent only one third of the possible maximum difference.

Furthermore, a particular situation is often having a large mean value across two or more social power bases. For example, this is true of situation 12 which reflects both reward and information bases of social power. Similarly, situation 10 reflects both expert and referent bases of social power. This might indicate that the six factor typology of social power is not as mutually exclusive as we are led to believe. It would be interesting to test this by examining the factor score correlations among the six uni-factors.

Finally, it is puzzling to note that the best fitting situations for each power base do not achieve comparable mean values. For example, while the mean values of the two situations depicting the expert base of social power is above 11, it is only 9.75 and 6.07 for the legiti-

mate base.

Overall, this paper is excellent. My only concern is that it has very little to do with consumer behavior, although it can be easily applied to study the process of persuasion and communication in marketing.

#### Directions for Future Research

While it is relatively easy to discuss and critique other people's research efforts, it is much harder to offer constructive suggestions as to where to go from here. There is no question that some time in the future, we will look back at each of the three studies as pioneering efforts in offering new directions in consumer behavior. I will attempt to offer some suggestions as to where the future research is likely to, or ought to emerge in each of the three areas.

#### Value Research

The biggest payoff in value research is likely to be in the development of consumption specific values, identifying their structure and creating a standardized psychometrically adequate scale. I will call this as the need to develop a Consumption Values Inventory (CVI). I think the real payoff in terms of explanation, prediction, dissemination and managerial use is likely to be somewhere between very fundamental values proposed by Rokeach and the multiattributes emphasized in attitude research. Furthermore, it should be specific, if not unique, to consumption behavior (Sheth, 1974). It is interesting to note that Daniel Yankolevich (1964) demonstrated the usefulness of consumption values for segmentation purposes at least a decade ago. Similarly, it is much more important and probably socially more relevant to examine consumer choice behavior at product class level rather than at brand choice level in linking consumption values to specific buying behavior. Due to the extreme market variety, predominantly artificial brand differences, and low degree of consumer involvement, it is probably neither possible nor relevant to build value based research at the level of brand choice behavior.

Third, we need to develop a comprehensive and large scale research on a representative national sample of the population to fully comprehend the missing linkages between a set of values and a set of product choice behaviors. For example, it is not very difficult to think of using the BLS typology of consumer expenditures and link them with consumption values. This will generate research findings which will be much more meaningful to both the public policy makers and the private marketing practitioners.

Finally, there is a considerable degree of opportunity to learn more about cross-cultural differences based on consumption values. This may explain more why the same product flourishes in one country and fizzles out in another. In many ways, Dichter (1962) presented some interesting cross-cultural anecdotes based on value research.

#### Deviant Consumer Behavior

The problem area of deviant consumer behavior needs several new research directions to be more substantive than issue-oriented research.

First, we need a much clearer definition of deviant consumer behavior which will offer a good typology of deviant consumer behavior by time, place and nature of deviancy. It is not enough to limit it to shoplifting, vandalism and deception in the retail store setting. This will be a very difficult task if the research on

consumer protection and consumerism is any guide as the other side of the coin.

Second, deviant consumer behavior is likely to be much more fruitful if the sociology of consumption tradition is brought to bear in theorizing about it rather than relying upon the psychological processes. This will inevitably force thinking and research at a more aggregate level in terms of groups or classes of consumers, and types of opportunity settings for deviant consumer behavior to manifest itself. It is my belief that any microlevel modeling at the individual consumer level will not be fully complete since there are too many time, place and person specific situational influences which intervene between a consumer's propensity for deviant behavior and its actual manifestation.

Finally, we need far more complicated theory than what has been proposed so far. As I mentioned earlier, deviant consumer behavior is more likely to be an interactive process. I have identified three factors: opportunity, consumer values possibly indicated by demographics, and perceived legitimacy of behavior in a particular setting. There can be more factors but the crucial point is that deviant consumer behavior is an interactive outcome of several factors.

#### Measuring Social Power

It is much more difficult to imagine the directions of future research related to social power. The concept has been already heavily utilized in terms of buyer-seller interaction process, opinion leadership and interpersonal communication.

I think the best way to bring social power theory more closely to consumer behavior is to drop the adjective "social" and focus more sharply on the role of power in a transaction or exchange between the buyer and the marketer. It would be most interesting to translate the traditional four elements of marketing mix (product, place, price and promotion) into the six bases of power (reward, coercion, legitimate, expert, referent, and information). In other words, is there a dominance of one or more power bases in each element of the marketing mix? Similarly, does the consumer have any power base, and if so, how does he acquire them?

A second area of research is to understand consumer perception of various power bases under different types of competitive structures in the market place. For example, a monopoly or even an oligopoly structure may be perceived as coercive power, regulated monopoly or perfect competition as legitimate power, service industry as expert or information power, and monopolistic competition as reward power. This may be most useful in assessing corporate image and the role of business institutions in society.

A third area of future research is to correlate consumer characteristics with preferences for specific power bases. For example, do higher socioeconomic people prefer information or expert base and lower socioeconomic people prefer reward or referent base of power? Do children prefer reward or coercive power bases more than adults, for example?

Finally, it will be most interesting to link specific products and services as means to exercising certain power bases. For example, food is often used for a reward base of power, appliances for a referent base, technical books for an expert base, mass media for an information base, and dangerous objects such as guns for the coercive base of power.

I hope these suggestions about the future directions of research in each of the three areas is useful to the

authors and others in consumer behavior.

#### References

- Gary A. Anderson and Jagdish N. Sheth, "A Comparative Analysis of Multidimensional Scaling of Portfolio Investment," Faculty Working Paper No. 344, October 13, 1976.
- Ernest Dichter, "The World Customer," Harvard Business Review, 40 (July - August, 1962), 116-120.
- John E. Overall and C. James Klett, Applied Multivariate Analysis, (New York, McGraw Hill, 1972), Chapter 17.
- Robert Redinger and Jagdish N. Sheth, "Multidimensional Scaling of Binary Data for Homogeneous Groups," Faculty Working Paper No. 411, June 29, 1977.
- Jagdish N. Sheth, "Problems and Potential of Multidimensional Scaling & Conjoint Analysis," Paper presented at AMA Educators Conference, Memphis, August 8-11, 1976.
- Jagdish N. Sheth, "The Next Decade of Buyer Behavior Theory and Research," in J. N. Sheth (ed.) Models of Buyer Behavior, (New York, Harper & Row, 1974), 391-406.
- Jagdish N. Sheth and Scott Armstrong, "Factor Analysis of Marketing Data: A Critical Evaluation," in P. R. McDonald (ed.) Marketing Involvement in Society and the Economy, (American Marketing Association, 1969), 137-144.
- Daniel Yankelovich, "New Criteria for Market Segmentation," Harvard Business Review, 42 (1964), 83-90.

## FAMILY COMMUNICATION AND CONSUMER SOCIALIZATION

George P. Moschis, Georgia State University  
Roy L. Moore, Georgia State University

### Abstract

This paper investigates family influences in terms of mediating socialization processes on adolescents' consumer skill acquisition. Family influences are first conceptualized as parent-child interpersonal processes, and hypotheses regarding their impact on the adolescent's consumer learning are formulated on the basis of theory and previous findings. Family influences are also assessed using path analysis.

### Introduction

In recent years, consumer learning among young people has become of increasing interest to several groups of people including marketers, public policy makers, consumer educators and students of socialization and consumer behavior (Ward 1974). In spite of the growing interest, relatively little research has been focused upon the process(es) by which young people acquire consumption-related skills, knowledge and attitudes--i.e., consumer socialization. Most of the published research in the area has focused upon television influences on consumer learning of children (e.g., Ward 1974) perhaps as a result of the recent issues surrounding the effects of marketing activities upon young consumers. Considerably less attention has been given to the examination of other sources of consumer influence (e.g., parents, peers) and age groups (e.g., adolescents) which are believed to be of great importance in the socialization of people to the consumer role (e.g., Moore and Stephens 1975, Moschis and Moore 1978, Ward et al 1977).

This paper investigates family influences on the acquisition of selected consumer skills during adolescence, a period believed to be crucial in socialization (Campbell 1969). Family influences are first conceptualized in terms of parent-child interpersonal processes. Hypotheses regarding the family's impact on adolescents' consumer socialization are then formulated on the basis of theory and previous research findings. Finally, the paper presents a path analytic model concerning family influences on adolescent consumer learning.

### Conceptualization of Family Influences

Past approaches to understanding how the family influences children ranged from very broad of global nature to situation specific. The first category includes structural variables relating to family characteristics such as social class, family size and ethnic background. Situational variables are more specific and "closer" related to the particular type of socialization such as parental consumer behavior and parent-child interactions about consumption (Ward et al 1977). Situational variables appear to have an advantage over global variables because they can often provide specific information regarding the influence.

One type of situation specific variable, parent-child communication about consumption, appears to be important in the socialization of children and adolescents (e.g., Ward et al 1977, Moschis 1976). However, little is known how specific communication patterns affect consumer socialization. Several studies in the related area of political socialization, for example, show that the spe-

cific pattern of parent-child communication has a more significant influence in socialization than parent-child frequency or amount of interaction (McLeod and Chaffee 1972).

Studies of general parent-child communication processes constantly find two relatively uncorrelated dimensions of communication structure: the first is called socio-oriented, the type of communication that is designed to produce deference and to foster harmonious and pleasant social relationships at home. The child in homes characterized by such a communication structure may be taught to avoid controversy and repress his feelings on extra-personal topics, for example, by not arguing with adults and giving in on arguments rather than risk offending others. The second type of communication is called concept-oriented, a pattern that focuses on positive constraints helping the child to develop his own views about the world. The parents may, for example, encourage the child to weigh all alternatives before making a decision or may expose him to controversy--either by differing openly on an issue or by discussing it with guests at home (McLeod and Chaffee 1972). The two general dimensions of parent-to-child communication produce a four-fold typology of family communication patterns (FCP): laissez-faire, protective, pluralistic, and consensual (McLeod and Chaffee 1972).

Laissez-faire families lack emphasis on either kind of communication; there is little parent-child communication in these families. Protective families stress obedience and social harmony in their communication with their child; there is little concern over conceptual matters. Pluralistic families encourage open communication and discussion of ideas without insisting on obedience to authority; the child is encouraged to explore new ideas and express them without fear of retaliation. The emphasis in this communication structure appears to be mutuality of respects and interests. Consensual families stress both types of communication; the child is encouraged to take an interest in the world of ideas, yet to do so without disturbing the family's hierarchy of opinion and internal harmony.

Extensive research evidence has led researchers to assume that these communication patterns help guide the child in coping with various situations he encounters outside the immediate family context--for instance, situations in relation to public affairs issues, school activity, and mass media use (e.g., Chaffee et al, 1966 and 1971; McLeod and Chaffee 1972); they have been extensively used in the area of political socialization and have predicted the person's learning in this area rather well (McLeod et al, 1968-1969; Chaffee et al 1970; Sheinkopf 1973). Since the area of political socialization is similar to consumer socialization in many ways (Ward 1974b), FCP may also predict consumer learning. The general hypothesis made in this study is that the adolescent's consumer behavior is conditioned by the structure of parent-child communication roles in the home. Specifically, the FCP as a generalized socializing influence would hypothetically lead to (a) different perception of consumer goods and services, and (b) different levels of competence of the adolescent's consumer skills.

## Hypotheses

Ward (1974b) speculated that families stressing conformity to others may implicitly encourage children to "learn to purchase and to derive satisfaction from their purchases on the basis of the perceived effects on others" (p. 40). Thus, a socio-oriented communication structure, which encourages the child to develop respect for others and other social orientations (Chaffee et al 1971) may lead to the development of materialistic orientations.

H<sub>1</sub>: Socio-oriented family communication structure is positively related to the adolescent's materialistic attitudes.

Because previous research found materialistic attitudes to be related to social utility motivations for watching television commercials and programs (e.g., watching commercials and programs to learn what products to buy to make good impressions on others) (Ward and Wackman 1971; Moschis and Churchill 1977), such motivations may also be the result of family communication structure at home. Thus, a socio-oriented family communication structure may implicitly encourage the child to pay attention to the mass media as a means of learning how to behave in various social settings.

H<sub>2</sub>: Socio-oriented family communication structure is positively related to the adolescent's social utility motivations for watching television commercials.

H<sub>3</sub>: Socio-oriented family communication structure is positively related to the adolescent's social utility motivations for watching television programs.

Research in the related area of political socialization revealed that a family communication environment stressing strong-concept orientations stimulates a greater political competence than an environment stressing socio-orientations. Furthermore, the pluralistic children tend to be more competent in political affairs than children from consensual homes because of the absence of social constraints in the former category. Specifically, pluralistic children were found to have greater political knowledge, were more sensitive to information contained in messages, and had greater ability to cognitively differentiate political stimuli (could point out similarities and differences among political candidates) than children from other types of communication environments (McLeod and Chaffee 1972). Similar findings may also apply in the area of consumer socialization.

H<sub>4</sub>: Adolescents from homes characterized by pluralistic family communication patterns (a) have greater consumer affairs knowledge, (b) are better able to filter puffery in advertising and (c) are better able to cognitively differentiate among products than adolescents from other types of family communication environments.

## The Study

The sample for this study consisted of 301 adolescent respondents attending junior and senior high schools of an urban and a semirural city in Wisconsin. Self-administered questionnaires were completed by students in various classes of three different schools.

### Family Communication Measures

Since the dependent variables were consumption related in nature, some of the traditionally used items to measure the two general parent-child communication structures

were revised to reflect communications more directly relevant to the consumer field. Items used to measure socio-orientation were:

(Parent) says the best way to stay out of trouble is to stay away from it.

(Parent) says his ideas are correct and (child) shouldn't question them.

(Parent) answers (child's) arguments with saying something like "You'll know better when you grow up."

(Parent) says (child) should give in when he argues rather than risk making people angry.

(Parent) tells (child) what things he should or shouldn't buy.

(Parent) wants to know what (child) does with his money.

(Parent) complains when he does not like something (child) bought for himself.

Responses to these items were measured on a 5-point "Very often-Never" scale. The concept-orientation relation was measured similarly and included the following items:

(Parent) says (child) should make his own decisions on things that affect him.

(Parent) emphasizes that every member of the family should have some say in family decisions.

(Parent) admits that children know more about some things than adults do.

(Parent) says that getting (child's) ideas across is important even if others don't like them.

(Parent) asks (child) what he thinks about things (parent) buys for himself.

(Parent) tells (child) why he buys some things for himself.

(Parent) tells (child) he should decide about things he should or shouldn't buy.

(Parent) tells (child) what he does with his money.

The reliability coefficient alpha of the two scales were .67 and .71, respectively, above the .50 to .60 reliability coefficients often recommended for constructs in the early stages of research (Nunnally, 1967, p. 226). The high internal consistency among general and specific items further suggests that the general family communication structures also apply to communication structures specifically related to consumption matters; and it provides validity for the revised items.

External validation for the revised items was also performed by correlating each family communication pattern with media use variables and comparing the results to those of previous studies that have used the traditional FCP measures but similar media use variables (e.g., Chaffee et al 1971). Table 1 shows that these correlations are fairly similar to those of previous studies. For example, adolescents in pluralistic homes give considerable attention to media news reports and spend relatively little time with television.

The sample for this study was divided into relatively "high" and "low" groups on each dimension by splitting each of the two scales at median, which yielded a four-



TABLE 1  
 ADOLESCENT'S MEDIA USE HABITS  
 BY FAMILY COMMUNICATION PATTERNS

Media Use Measures	Family Communication Pattern				Mean Raw Score Overall
	Laissez faire	Pluralistic	Protective	Consensual	
Amount of TV Viewing	-35*	-04	+14	+31*	23.14
TV Entertainment	-31*	+08	+04	+32*	19.77
TV News	-23*	+27*	-03	+05	3.37
News Reading	-13	+42*	-29*	-08	4.48
(Number of Cases)	(12)	(72)	(59)	(78)	(301)

Note: Entries are standard scores based on weighted means, setting the overall mean at zero and the standard deviation at unity within each row. Scores are calculated to decimal places; decimals are omitted for simplicity. The overall means for each dependent measure are shown at the right of the table. Asterisk(\*) indicates that group mean score is significantly different ( $p < .05$ ) from the combined mean score of the three remaining groups.

fold typology of family types with approximately equal numbers in each cell.

#### Criterion Variables

Consumer affairs knowledge referred to the accuracy of the cognitions held with respect to basic terms used in the marketplace as well as basic consumer-related legislation. This variable was measured by summing responses representing correct answers to 11 "true-false-don't know" items such as: "The mortgage is the down payment on a house" and "Milk sold in the store must show the last day it can be sold." The alpha reliability coefficient was .57.

Cognitive differentiation was operationally defined as the ability to identify products that are claimed to be different on specific attributes. This variable was measured by asking respondents to write the names of 12 products or brands selected at random during prime-time programs over a 3-month period and summing up items to form a 0- to 12-point index. A typical item was: "\_\_\_\_\_ camera weighs 16 ounces and costs \$66.00." The reliability coefficient was .86.

Puffery filtering referred to the respondent's ability to discriminate "facts" from exaggeration in advertising. Forty adult judges were presented with a long list of advertising claims containing various amounts of puffery and were asked to determine the degree of puffery in each item. The final list consisted of 12 items. Six of these items were considered (on the basis of the pretest) to contain the greatest amount of puffery; the other six were considered to contain "true" information. Respondents were given scores of 1, 2, and 3 for responses "believe it is completely true," "believe it is partly true," and "believe it is not true at all," respectively, given to those items that were considered to

contain the greatest amount of puffery; they were given scores of 3, 2, and 1 for providing similar responses to advertising claims considered to be true. Thus, the accuracy index could range from 12 to 36, with alpha reliability coefficient of .25.

Social Utility Reasons for watching television commercials and programs referred to adolescent motivations to watch TV commercials and programs as a means of gathering information about life styles and behaviors associated with uses of consumer products. Respondents were asked to indicate whether they ever watch television shows and commercials for ten different social reasons such as "to find out what qualities people like in others" and "to learn what things to buy to make good impressions on others." Responses were summed to form two 0- to 10-point scales: one which reflected social utility reasons for watching television commercials and another tapping the respondent's motivations for watching television programs, with alpha reliability coefficients of .64 and .70, respectively.

Materialism was operationally defined as "an orientation emphasizing possession and money for personal happiness and social progress" (Ward and Wackman 1971, p. 426). It was measured by soliciting responses on a 5-point "strongly agree-strongly disagree" Likert-type scale to six items, many of which were similar to Ward and Wackman (1971). The alpha reliability coefficient was .60.

#### Results

The first phase of data analysis dealt with the examination of the relationships between selected demographic characteristics of the respondents (age, sex, and social class) and measures of family communication structures and patterns. No significant relationship emerged between the three characteristics and the communication structures and patterns, suggesting that the extent of these communication processes among the respondents in the sample was not affected by their demographic characteristics. This finding eliminated the need for including age and other demographics as control variables in conducting further analysis.

The data supported Hypothesis 1. The correlation between socio-oriented family communication structure and materialism was statistically significant ( $r = .18$ ,  $p < .001$ ), while the relationship between concept-oriented communication structure and materialism was insignificant ( $r = .00$ ).

The resulting correlations between socio-oriented family communication structure and social utility motivations for watching television commercials was .21 ( $p < .001$ ), supporting Hypothesis 2. Similarly, the correlation between the socio-oriented communication structure and the respondent's motivations to watch television programs for social reasons was equally strong ( $r = .17$ ,  $p < .002$ ), which supports Hypothesis 3. These results suggest that families characterized by socio-orientation communication structure may be encouraging their children to turn to the media to learn appropriate social orientations or consumption behaviors appropriate to certain roles. This may in turn lead to the learning of materialistic orientations.

Table 2 shows mean values of measures of various consumer competencies by family communication pattern. As expected, students from pluralistic homes scored significantly higher on all three consumer competence measures. Children from such a family background seem to know more about consumer matters; they are better able to filter puffery in advertisements; and better able to

TABLE 2

ADOLESCENT'S CONSUMER SKILLS  
BY FAMILY COMMUNICATION PATTERNS

Consumer Skill Measures	Family Communication Pattern				Mean Raw Score Overall
	Laissez faire	Pluralistic	Protective	Consensual	
Consumer Affairs	-02	+26*	+08	-30*	6.30
Puffery Filtering	-08	+22*	+12	-20*	28.22
Cognitive Differentiation	-01	+28*	+03	-26*	6.33
(Number of Cases)	(92)	(72)	(59)	(78)	(301)

Note: Entries are standard scores based on weighted means, setting the overall mean at zero and the standard deviation at unity within each row. Scores are calculated to decimal places; decimals are omitted for simplicity. The overall means for each dependent measure are shown at the right of the table. Asterisk (\*) indicates that group mean score is significantly different ( $p < .05$ ) from the combined mean score of the three remaining groups.

cognitively differentiate among products as a result of their exposure to television commercials, even though pluralistic children do not watch significantly more television than their counterparts. These results support Hypothesis 4.

The final concern in this study was the investigation of the extent to which family communication structures directly or indirectly affect consumer learning. Direct impact on consumer learning may be the result of purposive parent-child interaction structures (e.g., purposive consumer training, acting as a consumer under parental guidance). Indirectly, the influence of family communication on consumer learning may be through influence of the child's interaction with other socialization agents, which in turn may influence learning. Previous researchers have reported that purposive consumer training is a rare occurrence at home (Ward 1974, Ward et al 1977), suggesting that the effect of family communication structure may indirectly affect consumer learning. Studies, in addition, have found that consumer and political skill acquisition strongly related to the child's frequency of reading and viewing news (consumer/political) (Chaffee et al 1971; Moschis 1976), which in turn relates to the family communication structures at home (Chaffee et al 1971). These findings suggest that family communication structures at home may indirectly affect the adolescent's acquisition of the three consumer competencies via public affairs media use.

On the basis of these speculations and in the context of a general conceptual model of consumer socialization discussed in greater detail elsewhere (e.g., Moore and Stephens 1975; Moschis 1976; Moschis and Moore 1978), path analysis was used to empirically address this question. Age and social class were included in the analysis as antecedent variables that may affect learning directly or indirectly (e.g., Ward 1974), while the two family communication structures and public affairs media use can

be viewed as intervening socialization variables (e.g., Moschis and Moore 1978).

Figures 1a, 1b, and 1c show results of this path analytic investigation, with the effects of antecedent variables and socialization processes on the three consumer learning skills reported in the form of path coefficients (betas). The results in all three figures suggest that the influence of the family communication structures on the learning of the three consumer skills examined may be indirect. Concept-oriented family communication structure apparently leads to differential exposure to the mass media, which in turn may lead to the learning of the three skills. The influence does not appear to be direct nor to be totally explained by the antecedent variables. Age and concept-oriented family communication structure are the best predictors of the adolescent's frequency of interacting with the media regarding public affairs content (frequency of reading news about the economy, government and advertisements; and viewing national and local news on television--measured on a 5-point "everyday-never" scale). Exposure to public affairs in turn seems to be the best predictor of all three skills examined.

#### Summary and Discussion

The present study investigated the effects of family influence on adolescent consumer learning in terms of mediating socializing processes. These processes were studied in the context of a two-dimensional model of communication which provides a four-fold typology of family communication patterns.

The results suggest that family communication structures, which appear to vary from family to family, may affect the child's perception of various marketing stimuli. Specifically, parents who emphasize the importance of social relationships (socio-oriented structure) in their communications with their children may implicitly encourage their children to evaluate their actions (including their consumer behaviors) on the basis of the perceived effects on others. This may result in the development of materialistic orientations. The findings also suggest that family communication structures at home as generalized socializing influences may lead to the differential exposure to mass media content, which in turn may result in learning of consumer competencies.

With respect to the influence of the various family communication patterns on the acquisition of the specific consumer competencies examined in this research, the data suggest that children from "pluralistic" families are more likely to have greater knowledge about consumer matters; they are better able to filter puffery in advertising; and they were better able to cognitively differentiate product-related information they are exposed to in the advertisements. These results regarding consumer competencies were similar to results regarding political competencies among adolescents.

#### References

- Earnest Q. Campbell, "Adolescent Socialization" in David A. Goslin (ed.) *Handbook of Socialization Theory and Research* (Chicago: Rand McNally, 1969).
- Steven H. Chaffee, Jack M. McLeod and Charles K. Atkin, "Parental Influences on Adolescent Media Use," *American Behavioral Scientist*, 14 (January-February, 1971), 323-340.
- Steven H. Chaffee, Jack M. McLeod, and Daniel B. Wackman, "Family Communication Patterns and Political Participation," Paper presented to the Association for Education

RESULTS OF PATH ANALYTIC INVESTIGATION

Figure 1a

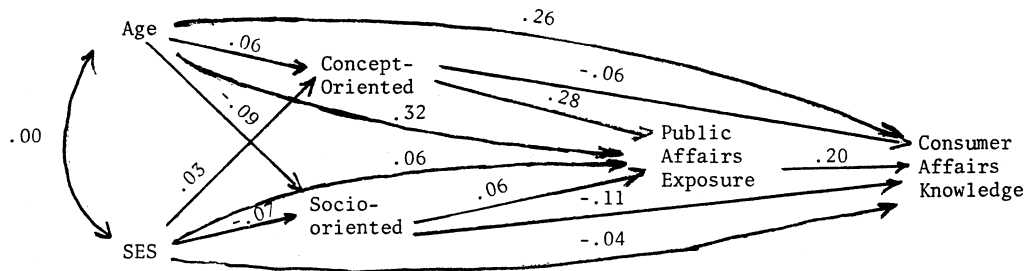


Figure 1b

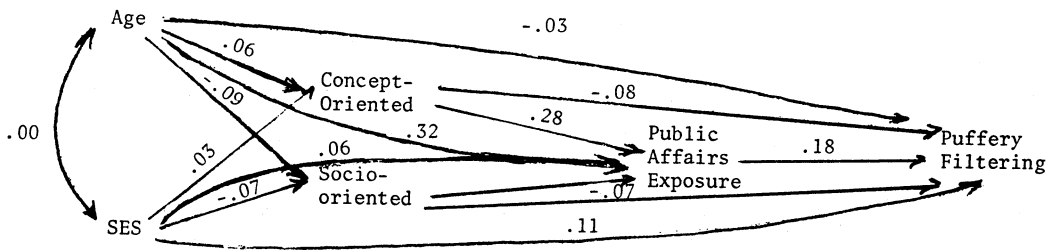
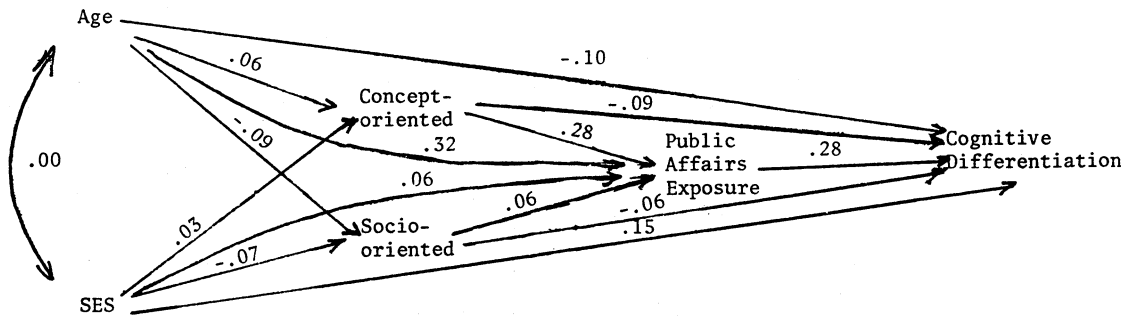


Figure 1c



NOTE: Path coefficients of approximately .11 or greater are significant at .05 level.

in Journalism (Iowa City: Iowa, 1966).

Steven H. Chaffee, Scott L. Ward, and Leonard P. Tipton, "Mass Communication and Political Socialization," Journalism Quarterly, 47(Winter, 1970), 647-659.

Jack M. McLeod and Steven H. Chaffee, "The Construction of Social Reality," in J. T. Tedeschi (ed.), The Social Influence Process (Chicago: Aldine-Atherton, 1972), 50-99.

Roy L. Moore and Lowndes F. Stephens, "Some Communication and Demographic Determinants of Adolescent Consumer Learning," Journal of Consumer Research, 2(September, 1975), 80-92.

George P. Moschis, "Acquisition of the Consumer Role by Adolescents," Unpublished Ph.D. Dissertation, University of Wisconsin-Madison, 1976.

George P. Moschis and Gilbert A. Churchill, Jr. "Mass Media and Interpersonal Influences on Adolescent Consumer Learning," Paper presented at the American Marketing Association Conference (Hartford, August 8-11, 1977).

George P. Moschis and Roy L. Moore, "An Analysis of the Acquisition of Some Consumer Competencies Among Adolescents," Journal of Consumer Affairs, forthcoming, 1978.

J. C. Nunnally, Psychometric Theory (New York: McGraw-Hill Company, Inc., 1967).

Scott Ward, "Consumer Socialization," Journal of Consumer Research, 1(September, 1974), 1-14.

Scott L. Ward, "Consumer Socialization," (Cambridge, Mass.: Marketing Science Institute, 1974b).

Scott Ward and Daniel Wackman, "Family and Media Influences on Adolescent Consumer Learning," American Behavioral Scientist, 14(January-February, 1971), 415-427.

Scott Ward, Daniel Wackman and Ellen Wartella, How Children Learn to Buy (Beverly Hills, Calif.: Sage Publications, 1977).

A CROSS-NATIONAL EXPLORATION OF HUSBAND-WIFE INVOLVEMENT<sup>1</sup>  
IN SELECTED HOUSEHOLD ACTIVITIES

Susan P. Douglas, Centre d'Enseignement Superieur des Affaires

Abstract

The paper presents the findings of an exploratory study of relative husband-wife involvement in a variety of household activities, based on data from couples in urban communities in five countries; two English speaking, - the U.S., the U.K.; and three French speaking, - France, Belgium and Canada. Similarities and differences in involvement for various types of activities between countries and between the two language groups are examined.

Introduction

Recent years have witnessed the emergence of a growing interest in studying household decision-making and task organization (Davis, 1976). This appears to have been stimulated largely by the radical changes which have been taking place in husband and wife roles, and the accompanying breakdown of traditional sex-role ideologies and cultural norms governing the division of labor and responsibility in the household. Concern has thus emerged in examining the impact and implications of such changes for consumer decision-making and behavior. Most studies of these issues have, however, been conducted by U.S. researchers, in relation to changes in U.S. society. Considerably less is known concerning the existence and impact of similar trends in other countries.

A number of studies of household decision-making and task involvement have been conducted by family sociologists in countries outside the U.S. These include studies in Belgium (Leplae, 1968; Silverman and Hill, 1967), Denmark (Kandel and Lesser, 1972), Finland (Haavio-Mannila, 1972), France (Michel, 1967), Germany (Lamouse, 1969; Lupri, 1969), Greece (Safilios-Rothschild, 1967, 1969), Japan (Blood, 1967), Mexico (Cromwell et al., 1973) and Yugoslavia (Buric and Zecevic, 1967). Inspired by the classic Blood-Wolfe (1960) study of the division of labor and responsibilities in households in the U.S., these are primarily concerned with testing the "resource" theory of marital roles.<sup>2</sup> They develop a measure or overall score of relative husband-wife influence for each household across a number of decisions and tasks, and examine the attitudinal and socio-economic determinants of this score in each country. Findings at this level of aggregation, i.e. across decisions and tasks, are, however, only indi-

rectly relevant for consumer behavior. Recent studies (Davis, 1976; Douglas and Wind, 1978), show, for example, that relative husband-wife involvement in a given household varies from one product category to another.

A few studies more specifically related to consumer decision-making in households outside the U.S. have also been made. Generally, these have been conducted in a single country, and with one exception (Hempel, 1974) there is little explicit attempt to make a cross-national comparison. Davis and Rigaux (1974) Rigaux-Bricmont (1977), for example, have investigated husband-wife involvement at different stages of the decision-making process for 25 economic decisions in Belgium. The Hempel study, on the other hand, examined husband-wife influence in relation to various stages and aspects of the decision to buy a house in both the U.S. and the U.K.

Comparison of findings in different countries is, however, hampered by lack of explicit hypotheses to guide interpretation, and to indicate where and why similarities or differences should be expected to occur. Furthermore, problems concerning the comparability and equivalence in each country, of a consumer decision such as the purchase of a house or a lawn mower also arise.

The purpose of the exploratory study reported in this paper was to gain some insights into the question of how to approach cross-national comparisons of household decision-making and consumer behavior. In particular, it was concerned with probing the basis on which countries should be selected, and with investigating which areas of family life were likely to be of interest for comparative purposes. Husband-wife involvement in a variety of household activities was, therefore, studied based on small samples of husbands and wives, living in urban areas in five different countries.

Since language group lines have been found in various cross-national studies to be a major parameter delineating differences and similarities in values and behavior (Haire et al, 1966; Hofstede, 1976; Roberts, 1970), the samples were selected from countries in two language groups - French-speaking and English-speaking in the North American and European continents. Language is also considered in cultural anthropology to be an important factor in defining appropriate units for comparison. One leading cultural anthropological methodologist has, for example, proposed the use of the culti-unit "a group bound by a common language, and belonging to the same state or territorial unit" in making cross-cultural comparisons<sup>3</sup> (Naroll, 1971).

<sup>1</sup>The data on which the study was based were collected in Belgium by Dr. Benny Rigaux, CESAM, in Canada by Professor Jean-Marie Lefebvre, Laval University, in the U.K. by Professor Peter Doyle, University of Bradford, and Professor Michael Baker, University of Strathclyde, and in the U.S. by Professor Harry Davis, University of Chicago. The author wishes to acknowledge their assistance as also the valuable aid of Professor Davis in the planning and analysis of the data. Financial support for analysis was provided by the Marketing Science Institute, Cambridge.

<sup>2</sup>The "resource" theory of marital roles hypothesizes that the relative power or influence wielded by the husband or the wife will be determined by the relative importance of the resources, e.g. education, income, they bring to the household. The principal competing theory is that of "ideological" role theory, which postulates that relative power or influence vs. determined by the dominant sex-role ideology, e.g. male-dominant, egalitarian in the family.

<sup>3</sup>While it is recognized that a country is not equivalent to a culture, and hence that a cross-cultural study is not necessarily cross-national, and v.v., one of the purposes here was to attempt to identify relevant dimensions in cross-national comparisons, of which culture (as operationalized in terms of language) might be one. There is also a substantial body of literature in socio-linguistics and anthropology, which postulates language to be an important factor in the formation of thought patterns, and modes of response, as in the development and transmission of behavior patterns (Cole and Scribner, 1974; Fishman, 1973; Hall, 1976; Hymes, 1967; Spier & al., 1941).

TABLE 1

## The Seventeen Household Activities

1. Taking care of savings and investments
2. Paying routine household bills
3. Balancing the checkbook
4. Taking out the garbage
5. Vacuuming the house
6. Calling in repairmen
7. Taking the car to the garage for servicing
8. Making travel reservations for the family
9. Going shopping for furniture
10. Suggesting going out for dinner
11. Deciding to have the husband's suit cleaned
12. Buying records
13. Going shopping for a new car
14. Deciding what clothes the husband should wear when going out to dinner
15. Going to the supermarket
16. Taking care of the garden
17. Giving the children a bath

The household activities examined, covered a broad spectrum of family life. They ranged from traditionally sex-specialized tasks and decisions such as giving the children a bath, going to the supermarket (wife-dominated), taking the car for servicing (husband-dominated), and traditionally joint or shared decisions and tasks, such as going shopping for furniture, to areas where cultural role expectations are less clear cut, such as deciding to have the husband's suit cleaned, calling in repairmen, or inviting guests for dinner.

The data base of the study and the research methodology are first presented. Then, the findings concerning the relative involvement of husbands and wives in the various activities are studied, and similarities and differences between the five groups examined. The implications of these findings are discussed, and some suggestions and directions for future research outlined.

## The Research Approach and Plan of Analysis

## Sample Characteristics and Data Base

The data base consisted of questionnaires administered separately to couples living in major urban areas in each country. These were Chicago, London, Glasgow, Paris, Brussels and Quebec. Due to the cost and difficulty of obtaining responses from both husbands and wives, convenience sampling was used.<sup>4</sup> The samples are somewhat "upscale" in all cases except for Quebec (see Appendix 1). The same questionnaire was filled out by both spouses, and couples were instructed not to consult each other during the completion process.

Respondents were asked questions relating to involvement in decisions and tasks, for two types of situations, one general, and the other specific: 1) who was typically responsible for each of 17 general household activities (as listed in Table 1) and 2) who was responsible for performing each of six tasks the last time guests were invited for dinner.<sup>5</sup> In addition, a number of questions were asked relating to family background characteristics, including standard variables such as income, education of husband and wife, as well as descriptors of their home, such as ownership of household appliances, number of rooms and type of dwelling.

<sup>4</sup>The sample size is small in each urban area, ranging from 160 couples in Chicago to 76 in Brussels, 95 in Quebec, and 101 in Paris. The use of small sample sizes is, however, by no means uncommon in family research (Starch, 1958; Davis, 1970; Davis and Rigaux, 1974) particularly when responses are obtained from both husband and wife, since costs and administrative difficulties are substantially increased. In this particular study, typically less than one-quarter to one-third of the questionnaires initially distributed, were returned or satisfactorily completed by both husband and wife.

<sup>5</sup>Initially following the classification of marital role categories identified by Herbst (1952), five response categories were used: "mainly husband" (husband dominant); "mainly wife", (wife-dominant); "sometimes husband, sometimes wife" (autonomic) and "generally both together" (syncratic) and "other". In the analysis, however, the autonomic "sometimes husband, sometimes wife" and syncratic "generally both", categories were combined, as there appeared to be considerable ambiguity concerning this distinction for respondents. The "other" category was also eliminated from the computations since the number of responses was minimal in all samples.

## Plan of Analysis

Husband-wife involvement in the two types of situations were analyzed separately. First, relative involvement in the seventeen household activities was compared. Similarities in the distribution of responses for each activity between five urban groups were examined. Then, following Wolfe (1959) two separate dimensions of husband-wife involvement were distinguished; the relative involvement of the husband as opposed to the wife and the proportion of couples who shared responsibility for an activity. These dimensions were examined overall, i.e. across all seventeen activities, and by each individual activity.<sup>6</sup>

Spearman correlations of rank order on the 17 activities were calculated between all possible pairs of the five groups, (i.e. ten in all) for each of the two dimensions of husband-wife involvement. McQuitty hierarchical clusterings for these two sets of correlations were then performed (McQuitty, 1960). Next, where differences between the five groups for an activity occurred on either dimension, these were studied in more detail, and differences within and between the two language groups examined. Similar procedures were used to examine task performance in the dinner situation. The same two dimensions of involvement were identified, and used to make comparisons between the five groups. The results of this analysis did not, however, add significantly to those of the general household situations, beyond confirming certain tendencies and hence are not discussed here in more detail.

Finally, the relation of husband-wife involvement to certain socio-demographic characteristics was examined. Not only, according to the "resource" theory of marital roles are these key determinants of the allocation of household responsibilities (Blood and Wolfe, 1960) but they have been found in various studies to be related to roles in consumer purchasing decisions (Davis, 1972; David, 1976; Green and Cunningham, 1975; Hempel, 1975). In addition, previous studies of marital roles and authority in various countries have revealed some differences in the strength of these relationships (Davis, 1972; Michel, 1967; Silverman and Hill, 1967). In some

<sup>6</sup>Responses of husbands and wives were pooled in the analysis. Some incongruence in responses occurred and hence it was considered preferable at this aggregate level to include both responses in order to avoid a "sex" bias.

cases the way in which these are related to the balance of authority in the family has been found to differ, but in countries at lower levels of economic development than those examined in this study.

Five background characteristics previously found to be related to marital roles - husband's education, wife's employment status, family income, type of dwelling and husband's age, were selected for examination. These were cross-tabulated with the two measures of husband-wife involvement for each of the seventeen activities to see whether observed similarities and differences might be accounted for, or related to differences in background characteristics between the samples, and lambda asymmetric statistics calculated to test the degree of association.<sup>7</sup>

### Research Findings

The initial inspection of husband-wife involvement for the 17 household activities showed a substantial degree of similarity between the five sample groups. As shown in Table 2, for nine of the activities, the degree of husband-wife involvement was much the same for all five groups. Whether in Chicago, Glasgow, Paris, Quebec or Brussels, husbands are mainly responsible for taking the car to the garage for servicing, wives with vacuuming and bathing the children, and both are involved in going shopping for furniture, and deciding to go out in the evening. Equally, going shopping for a new car is either a husband-dominated or joint activity, while involvement in paying bills, buying records and making travel reservations varies from one household to another.

TABLE 2

Similarities in Husband-Wife Involvement in the Five Urban Groups

Activity <sup>b</sup>		Percentage of respondents indicating each involvement category <sup>a</sup>				
		Chi-cago	London/ Glas-gow	Paris	Brus-sels	Que-bec
<b>Husband-Dominated</b>						
Taking the car to the garage for servicing	H	62	63	75	66	85
	W	14	10	8	3	3
	J	24	27	17	31	12
	Base	312	152	208	144	170
<b>Wife-Dominated</b>						
Vacuuming the house	H	4	4	11	2	7
	W	80	79	66	64	66
	J	16	17	23	34	27
	Base	291	164	172	131	176
Giving the children a bath	H	4	3	-	-	1
	W	63	69	84	73	24
	J	23	28	16	27	75
	Base	188	99	142	80	113
<b>Joint</b>						
Going shopping for furniture	H	1	1	5	3	2
	W	18	12	11	15	7
	J	81	87	84	82	91
	Base	322	164	209	148	80

<sup>7</sup>The lambda asymmetric statistic tests the degree of association for the case in which the dependent variable is nominally scaled, and the independent variable is nominally (or ordinally) scaled. No assumptions of normal distribution of either variable are made.

Activity <sup>b</sup>		London/ Chi-cago Glas-gow Paris Brus-sels Que-bec				
		Chi-cago	Glas-gow	Paris	Brus-sels	Que-bec
Suggesting going out for the evening	H	12	16	19	19	12
	W	13	17	19	20	17
	J	75	67	62	61	71
	Base	318	169	206	145	179
<b>Mixed</b>						
Paying household bills	H	42	48	34	33	45
	W	42	23	36	34	29
	J	16	29	30	33	26
	Base	319	171	216	151	181
Going shopping for a new car	H	34	41	48	43	43
	W	2	3	2	2	3
	J	64	56	50	55	54
	Base	307	148	205	135	164
Buying records	H	27	36	28	26	33
	W	22	12	17	15	22
	J	51	52	55	59	45
	Base	241	132	170	110	132
Travel reservations	H	57	55	40	52	49
	W	16	16	29	16	13
	J	27	29	31	32	38
	Base	304	152	197	136	158

<sup>a</sup>The three involvement categories were mainly husband, (H), mainly wife (W) and sometimes husband, sometimes wife or generally both together (J).

<sup>b</sup>The activities shown for the husband-dominated wife-dominated and joint categories included all those where at least 60% of respondents in all five groups indicated a dominant category. The activities shown for the "mixed" activities included all those where the range between the five groups was less than 20% for all three categories.

A more detailed examination of husband-wife participation in the seventeen activities, separating the two dimensions of shared and husband involvement reveals, however, a number of differences between the five groups. Both at the overall and at the individual activity level, these differences occur primarily between the two language groups.

### Overall Comparison of Shared and Husband Involvement Between the Five Groups

The McQuitty clustering of the correlations of the shared involvement measures shows that the nature or proportion of joint activity across all seventeen activities was most similar between the Chicago and London/Glasgow pair (Figure 1). The correlation coefficient for this pair was .91. The next most similar were Paris and Quebec with a coefficient of .89, joined by Brussels at the .78 level. The three French-speaking samples were thus more similar to each other than to the Chicago/London-Glasgow pair.

In the case of husband involvement the closest pair were Paris and Brussels, both characterized by similar levels of husband involvement across all activities (correlation coefficient .95) (Figure 2). Quebec again, was more similar to this pair than to Chicago or London-Glasgow (joining them at the .91 level). Thus, again, similarities within the two language groupings emerge, though in this case, less marked than in relation to shared involvement.

TABLE 4

## Relation Between Selected Background Characteristics and Participation in Activities for the Five Samples

Lambda asymmetric statistics for the background characteristics <sup>a</sup>					
Activity	Husband's Age	Husband's Educ.	Working Wife	Income	Apt
Joint					
Savings and Investment	Paris (.11) Brussels (.12)		Paris (.14)	Paris (.11) Brussels (.19)	
Checkbook					
Supermarket	Quebec (.15)		Paris (.13)		
Husband dominated					
Garbage		Glas./London (.11) Chicago	Glas./London (.20) Chicago (.15)	Brussels (.40) Glas./London (.13) Chicago (.19)	
Repairmen	Brussels (.44) Glas./London (.18) Chicago (.13)	Paris (.16) Glas./London (.17) Quebec (.12)	Brussels (.23) Glas./London (.14)	Paris (.13) Brussels (.10) Quebec (.13) Glas./London (.24)	Glas./London (.24)
Suit	Glas./London (.14)	Glas./London (.11)		Glas./London (.22)	
Clothes				Paris (.14)	
Garden	Chicago (.26)	Chicago (.11)		Brussels (.12) Chicago (.13)	

<sup>a</sup>lambda asymmetric statistics over .10

in the French language samples the wife had a more significant role. Equally, husbands in the latter group were more involved in taking out the garbage than in the former group. In doing the garden and calling in repairmen, however, the most marked differences occurred between Chicago and the other samples. In both cases, husband involvement was substantially lower in Chicago than elsewhere.

#### Relation Between Husband-Wife Involvement and Background Characteristics

Examination of the relation between the five background characteristics and measures of husband-wife involvement showed few systematic relationships of a nature which might account for the similarities and differences observed between the five groups. As Table 4 shows, relatively few of lambda statistics were over .10.<sup>8</sup> Husband's age was related to husband involvement in calling in repairmen in Brussels, and in taking care of the garden in Chicago. The wife's employment status was related to husband involvement in taking out the garbage in Glasgow and London, and in calling in repairmen in Brussels. Family income was related to husband involvement in taking out the garbage in Brussels, in calling in repairmen and in taking the husband's suit to the cleaners in London and Glasgow.

Since, however, the lambda statistic is a relatively rigorous test of association, and in view of findings of significant relationships in previous studies, a further check was provided by breaking down each sample into comparable subgroups. The degree of shared and husband involvement in working wife, and non-working wife families, in young and old couples, low and high education couples, low and high income couples, and apartment and house-dwellers, in the five samples was then compared. Invariably, the same tendencies and differences observed between the total samples occurred. Husbands in working wife families, for example, in Chicago and London-Glasgow were more involved in taking their suit to the cleaners, and were less likely to participate in grocery shopping than in working wife families in Paris, Brussels, and Quebec. Similarly, whether in high or low income families, there was less joint decision-making about savings and investment in Chicago and London-Glasgow than in Paris, Brussels and Quebec. In a few instances, there were some differences, but these occurred in cases where sample sizes were small, for example, taking out the garbage, or calling in repairmen, and hence, are somewhat ambivalent.

Consequently, while household socio-economic characteristics affected the degree of shared or joint involvement, they appear to do so in a similar way in each sample, and are thus unlikely to account for the major differences observed between the samples.

#### Discussion and Conclusions

The study thus provides a number of interesting indications concerning husband-wife involvement in the five samples. These should, however, be regarded essentially as providing guidelines for further research, particular-

<sup>8</sup>Calculation of gamma statistics, an ordinal symmetric test, or Yule's Q for the 2 x 2 tables, resulted in a substantially larger number of significant relationships ranging from 20 out of 40 cross-tabulations for type of dwelling, to 33 out of 40 for income. The lambda statistic is, however, a more appropriate test, given the nature of the variables.

ly in view of the sample sizes and sampling procedures. Further investigation of the various points on larger samples is required before any conclusions can be drawn concerning similarities and differences between the five groups. The key points arising from the study are next discussed, and directions for further research outlined.

#### Key Findings

First, it appears that with regard to a number of activities husband-wife involvement is substantially similar in all five samples. This is consistent with the conclusions of family sociologists who have found many similarities in involvement in various household decisions, and to a less extent in relation to husband-wife participation in household tasks (Blood and Hill, 1970; Silverman and Hill, 1967). It is also consistent with the findings of a comparative study of housing purchase decisions in the U.S. and the U.K. (Hempel, 1974).

The study extends these conclusions to cover a wider range of family activities, and suggests that similarities occur not only in relation to activities traditionally central to husband and wife roles as, for example, the car, vacuuming, the children, and activities involving both, such as shopping for furniture or going out for dinner, but also those where husband and wife involvement varies from household to household, such as making travel reservations and buying records. Consequently, findings of similarities do not appear to arise from lop-sided focus on areas of family life in relation to which cultural role expectations are traditionally clearly defined, and are dominated by husband or wife, but occur throughout the broad spectrum of the way in which husbands and wives carve up household responsibilities.

Secondly, it appears that where differences in relative involvement do occur between the five samples, they frequently divide up along language group lines. This is particularly noticeable in relation to two types of activities - those relating to husband's clothing, and certain household maintenance activities. In the latter case, these appear to be activities which traditionally have been the responsibility of one or the other spouse, but in which current trends differ from one language group to another.

In the case of husband's clothing, the greater degree of wife involvement in the French language group suggest that she has a more important role and influence in relation to male clothing decisions. One possible explanation for this finding is that in Latin countries, clothing and fashion are perceived as essentially the woman's domain, leading to greater involvement of the wife in male clothing and care decisions (Boss, 1976; Douglas, 1976). In Anglo-Saxon countries, there is greater male independence with regard to clothing and personal grooming decisions, and hence less female involvement.

Similarly, high levels of shared or husband involvement in traditionally wife-specialized tasks such as going to the supermarket, vacuuming, and taking out the garbage, suggest a greater crossing of the line, and aid given to wives by husbands in the French language groups. Equally, there was greater joint involvement in the French language group in the otherwise husband-dominated activity of savings and investment. Similar differences did not, however, occur in relation to another traditionally male dominated area - the car. The findings are somewhat contradictory to those of an earlier study comparing husband-wife task involvement in the U.S. and Belgium (Silverman and Hill, 1967) in which U.S. husbands were observed to participate more than Belgian husbands in household tasks. This may, however, be explained in part by the time lapse between the two studies and by the evolution of marital and sex role ideologies in the two language groups.

Even prior to the current trend towards equality between the sexes, egalitarian role norms have generally been more prevalent in Anglo-Saxon than in Latin cultures (Rocheblave-Spenle, 1964). The advent of "women's lib" has reinforced such tendencies, and opened up areas predominantly the prerogative of one sex to the other. How this actually affects relative husband-wife involvement appears to depend on the specific activity concerned. In the U.S., for example, while more joint involvement has taken place in automobile purchasing, traditionally a male bastion, over the last eighteen years, the purchase of insurance has become even more husband-dominated, and shopping for groceries more wife-dominated (Cunningham and Green, 1974).

In Latin cultures, on the other hand, acceptance of egalitarian role norms is more recent. Consequently, households mark this new development by husband involvement in traditionally female tasks, and joint activity in certain traditionally male-dominated areas. Findings of a similar nature have been noted in the family sociology studies of marital authority. These show that in egalitarian cultures, families with highly educated husbands tend to be male-dominated, while in patriarchal cultures, husbands with high educational levels cede authority to their wives. This latter finding is interpreted as a demonstration by the husband of his more advanced and liberated views relative to prevalent cultural role norms (Rodman, 1967).

#### Directions for Future Research

From these exploratory findings, four major areas for future research can be identified:

1. In the first place, the somewhat complex way in which husband and wives are involved in different areas of family life in the five community groups, suggests the need to compare decision and task involvement across a broad range of activities. With one or two exceptions, few activities seem to be systematically categorized across all five groups as husband-dominated, wife-dominated or joint. Different views about who should do what could thus underlie observed differences and similarities in relative involvement. In this study, for example, whether clothing is perceived as the wife's domain, and savings decisions as requiring joint discussion seem likely to be important parameters determining relative involvement.

Future research should therefore identify for each cultural group which activities are perceived as primarily the responsibility of one or the other spouse, or in which both should be involved, as well as which are perceived as central to family life and household organization. Involvement in all these activities could then be compared across all cultural groups to see to what extent similarities or differences were related to such perceptions.

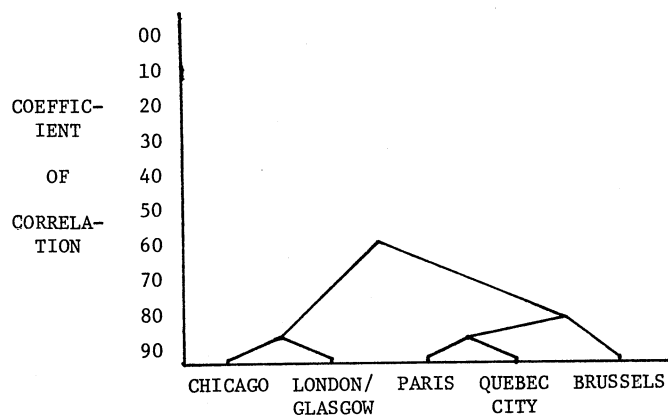
2. The differences in relative involvement observed between the two language groups suggest that this is a useful unit of analysis for future studies. A more extensive comparison of relative husband-wife involvement in English and French-speaking countries, based on nationally representative samples would appear a logical next step. Extension to other major language groups, such as a pair of Spanish language countries, could also be undertaken. Such research should focus on investigating the extent to which differences between language groups are greater than differences within language groups, and hence, provide an appropriate consumer grouping for market segmentation.

The impact of language as opposed to other macro-cultural variables, such as religion or the importance of the family as an institution, in determining relative husband-wife involvement could also be examined. For example,



FIGURE 1

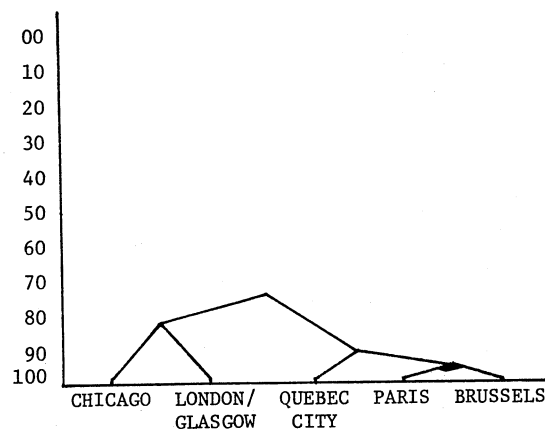
The Five Groups on the Shared McQuitty Clustering of Involvement Measures<sup>a</sup>



<sup>a</sup>The shared involvement measure was calculated on the basis of the proportion of "sometimes husband, sometimes wife" and "both" responses relative to the total response for the activity.

FIGURE 2

McQuitty Clustering of the Five Groups Based on the Degree of Husband Involvement Measure<sup>a</sup>



<sup>a</sup>The degree of husband involvement measure was calculated on the basis of the proportion of "mainly husband" responses relative to the total of both "mainly husband", and "mainly wife" responses.

Examination of Differences Between the Five Groups in the Degree of Shared and Husband Involvement in Various Activities

A more detailed examination of similarities and differences between the five groups by each activity showed that, for the measure of shared involvement, substantial differences (i.e. range between lowest and highest sample of over 20%) occurred for only three activities - taking care of savings and investment, balancing the checkbook, and going to the supermarket. As Table 3 shows, in two cases, savings and investment, and going to the supermarket, the differences occurred predominantly between the two language groups. In the English-language group

the degree of shared involvement was lower, averaging .35 and .25 for these two activities, as compared with .57 and .42 for the French-language group. In the case of the checkbook, however, Chicago was at one end of the spectrum with 12% shared involvement, and Quebec at the other, with 45%, while the three European samples ranged between 20 - 30%.

In regard to the degree of husband dominance, differences occurred in relation to five activities, and again, predominantly between the two language groups. In the two English language samples, the husband was primarily responsible for taking his suit to the cleaners, and for deciding what clothes he should wear for dinner, while

TABLE 3

Differences in the Degree of Shared and Husband Involvement in the Five Groups

Activity <sup>a</sup>	Percentage of respondents indicating shared or husband dominated involvement						
	Chicago	London Glasgow	English group	Paris	Brussels	Quebec	French Group
<b>Shared</b>							
Savings and Investment	34	36	35	48	54	68	57
Checkbook	12	27	19	20	30	36	29
Supermarket	22	26	24	38	43	45	42
<b>Husband-Dominated</b>							
Suit cleaned	72	56	64	20	16	21	19
Clothes for dinner	83	68	75	58	42	60	53
Garbage	57	50	54	58	63	80	70
Garden	49	71	55	80	68	74	74
Repairmen	37	50	43	57	52	67	59

<sup>a</sup>The activities considered different were those where the range between the highest and lowest sample was over .20, for the shared involvement measure, and over .30 for the husband-dominated measure.

Catholic or Protestant families in English and French language countries could be compared. Equally, comparisons of regions or urban vs. rural families within and between the major language groups of countries, could be made.

3. A systematic investigation of the relationship between household socio-economic characteristics and husband-wife involvement is also an important research priority. In this regard, two issues can be distinguished. First, since socio-economic or demographic characteristics are frequently used to draw national samples, differences between national samples may reflect differences in the distribution of these variables from country to country, rather than in the dependent variable. Preliminary investigation of this issue did not suggest this to be a major problem here. More systematic and detailed investigation is, however, clearly required.

Secondly, socio-economic and demographic characteristics such as the wife's employment, or the husband's education are frequently hypothesized to be important determinants of relative husband-wife involvement (Heer, 1963; Hoffman, 1960). Comparison of the impact of such variables from country to country gives rise, however, to a number of questions concerning their cross-national comparability and equivalence. Income levels relate to different standards and costs of living, occupations differ in their earning power, social position or prestige in different countries, and even family life-cycle is affected by socio-economic factors, since young adults in wealthy societies can afford to marry younger and live longer than in poorer societies (Blood and Hill, 1970). One approach to reduce such problems is to construct indicators of equivalent or comparable constructs, rather than directly comparing income or education variables. Indicators of social status might, for example, be developed by combining various education, income and occupation variables or of the wife's time pressure from the presence and number of young children and her occupation status (Davis, 1976).

4. Cultural role norms and the attitudes of individual family members towards such norms constitute another set of variables which need to be examined. It was earlier hypothesized that differences in the prevalence and tradition of egalitarian role norms in the two language groups might account for some of the observed differences in relative involvement. More explicit investigation and testing of this hypothesis is required, examining husband-wife involvement in countries or environments covering a wider range of dominant cultural norms or traditions.

Examination of the extent to which individual households subscribe to prevalent role norms may also aid in understanding how they divide up responsibility for various decisions and tasks. Since relative involvement varies from one household activity to another, and not along any simple clearcut lines, the impact of these cultural role norms needs to be examined relative to specific tasks and activities. Consequently, in addition to general attitudes to husband dominance, or egalitarian sex roles, attitudes towards involvement in specific tasks and decisions, as for example, husband involvement in household chores, or wife participation in financial decisions should also be investigated.

The study thus provides a number of guidelines for future cross-national comparisons of the division of labor and responsibility in the household. In particular, it suggests key types and categories of activities to examine, as well as appropriate units of analysis for making such comparisons. If the list of unresolved questions, and methodological problems in such research appears somewhat formidable, the potential insights to be gained not only concerning household task involvement in other countries, but also in understanding the basic forces at work and

underlying husband-wife interaction in the household, should encourage continued research and investigation.

## APPENDIX

### APPENDIX 1 - SAMPLE BACKGROUND CHARACTERISTICS

	Chi- cago	Lon- don/ Glas- gow	Paris	Brus- sels	Que- bec
<u>Husband's Age:</u>					
a. 30 or less	28	34	33	55	12
b. 31-40	34	36	27	15	16
c. 41-50	27	23	24	7	36
d. Over 50	11	8	16	23	36
<u>Husband's Education:</u>					
a. Grade School (9yrs)	1	6	9	-	} 27
b. Some high school (9-12yrs)	13	8	23	7	
c. High School (13-16yrs)	8	13	26	31	36
d. College (17-20yrs)	57	37	30	49	14
e. Postgraduate (20+)	30	36	12	13	23
<u>Total Family Income:</u>					
a. Less than \$5,00	-	7	4	18	10
b. \$5,000-\$10,000	2	26	15	38	29
c. \$10,000-\$15,000	14	36	23	21	27
d. \$15,000-\$20,000	26	17	19	13	20
e. \$20,000-\$25,000	20	7	15	4	6
f. Over \$25,000	38	7	24	6	8
<u>Working Wives</u>	46	57	56	50	23
<u>Apartment</u>	29	21	68	46	36

## REFERENCES

- Robert O. Blood and Robert L. Hamblin, "The Effects of Wife's Employment on the Family Power Structure", Social Forces, Vol. 36, May 1958, pp. 347-352.
- Robert Oscar Blood Jr., and Donald M. Wolfe, Husbands and Wives (Glencoe, Ill.: The Free Press, 1960).
- Robert O. Blood, Love Match and Arranged Marriage (New York: The Free Press, 1967).
- Robert O. Blood and Reuben Hill, "Comparative Analysis of Family Power Structure - Problems of Measurement and Interpretation", in Reuben Hill and Rene Konig (eds.) Families in East and West (Mouton: Paris, 1970).
- Jean-Francois Boss, "Elements de Qualification du Vêtement Masculin: Analyse Secondaires", Cahier de Recherche no. 53 bis, CESA, Jouy-en-Josas, 1976.
- Olivera Buric and Andjelica Zecevic, "Family Authority, Marital Satisfaction and Social Network in Yugoslavia", Journal of Marriage and the Family, Vol. 29, no. 2 (May 1967) pp. 325-336.
- Michael Cole and Sylvia Scribner, Culture and Thought (John Wiley & Sons: New York, 1974)
- Ronald E. Cromwell, Ramon Corrales and Peter M. Torsiello, "Normative Patterns of Marital Decision-making Power and Influence in Mexico and the United States: A Partial Test of Resource and Ideology Theory", Journal of Comparative Studies, Vol. 4 (Autumn, 1973) pp. 175-196.

- Harry L. Davis, "Dimensions of Marital Roles in Consumer Decision-making", Journal of Marketing Research, Vol. 7, 1970, pp. 168-177.
- Harry L. Davis, "Determinants of Marital Roles in a Consumer Purchase Decision", Working Paper 72-14, European Institute for Advanced Studies in Management, Brussels, April 1972.
- Harry L. Davis and Benny P. Rigaux, "Perceptions of Marital Roles in Decision Processes", Journal of Marketing Research, Vol. 8 (August 1971) pp. 51-61.
- Harry L. Davis, "Decision-making within the Household", Journal of Consumer Research, Vol 2, no. 4 (March 1976), pp. 241-260.
- Susan Douglas, "Men's Leisurewear: A Male-dominated Market", European Journal of Marketing, Vol. 1, no. 10 (Spring 1976).
- Susan Douglas and Yoram Wind, "Examining Family Role and Authority Patterns: Two Methodological Issues", Journal of Marriage and the Family, forthcoming February 1978.
- Joshua Fishman, Sociolinguistics (Rowley, Mass: Newbury House, 1973).
- Robert T. Green and Isabella Cunningham, "Feminine Role Perception and Family Purchasing Decisions", Journal of Marketing Research, Vol. 12 (August 1975) pp. 325-332.
- Mason Haire, E.E. Ghiselli and L. W. Porter, "Cultural Patterns in the Role of the Manager", Industrial Relations, Vol. 2 (1963) pp. 95-117.
- Edward T. Hall, Beyond Culture (New York: Anchor Press, 1976).
- Elina Haavio-Mannila, "Some Consequences of Woman's Emancipation", American Sociological Review, Vol. 25 (April 1970) pp. 161-178.
- David M. Heer, "The Measurement and Bases of Family Power", Marriage and Family Living, Vol. 25, no. 2 (May 1963) pp. 133-139.
- Donald J. Hempel, "Family Buying Decisions: A Cross-Cultural Perspective", Journal of Marketing Research, 11 (August 1974) pp. 295-302.
- Donald J. Hempel, "Family Role Structure and Housing Decisions", Advances in Consumer Research, Vol. 2 (ed. Mary Jane Schlinger) (Association for Consumer Research: Chicago, 1975) pp. 71-80.
- P. G. Herbst, "The Measurement of Family Relationship", Human Relations, Vol. 5 (February 1952) pp. 3-35.
- Lois Hoffman, "Effects of Employment of Mothers on Parental Power Relations and the Division of Household Tasks", Marriage and Family Living, Vol. 22, (February 1960) pp. 27-35.
- Geert Hofstede, "Nationality and Espoused Values of Managers", Journal of Applied Psychology, Vol. 61, no. 2 (1976) pp. 148-155.
- Dell Hymes, "The Anthropology of Communication", in Frank E. X. Dance (ed.), Human Communication Theory (New York: Holt, Rinehart and Winston, 1967).
- Denise B. Kandel and Gerald S. Lesser, "Marital Decision-making in American and Danish Urban Families", Journal of Marriage and the Family, Vol. 34, no. 1 (February 1972) pp. 134-138.
- Annette Lamouse, "Family Roles of Women", Journal of Marriage and the Family, Vol. 31, no. 1 (February 1969) pp. 145-152.
- Claire Le Plae, "Structure des Taches Domestiques et du Pouvoir de la Decision de la Dyade Conjugale", Pierre de Brie (ed.) La Vie Conjugale (Bruxelles: Editions Vie Ouvriere, 1968).
- Eugene Lupri, "Contemporary Authority Patterns in the West German Family", Journal of Marriage and the Family, Vol. 31, no. 1 (February 1969), pp. 134-144.
- Louis L. McQuitty, "Hierarchical Syndrome Analysis", Educational and Psychological Measurement, Vol. 20, no. 2 (Summer 1960) pp. 293-304.
- Andree Michel, "Comparative Data Concerning the Interaction in French and American Families", Journal of Marriage and the Family, Vol. 29, no. 2 (May 1967), pp. 320-324.
- Raoul Naroll, "Some Thoughts on Comparative Method in Cultural Anthropology", in Hubert Blalock and Anne Blalock; Methodology in Social Research, (McGraw Hill: International Student Edition, 1971).
- Benny Rigaux-Bricmont, La Prise de Decisions Economiques dans la Famille (Louvain: CESAM 1977).
- Anne-Marie Rocheblave-Spenle, Les Roles Masculins et Feminins (Paris: Presses Universitaires de France 1964).
- Hyman Rodman, "Marital Power in France, Germany, Yugoslavia, and the U.S.", Journal of Marriage and the Family, Vol. 29 (May 1967) pp. 320-324.
- Hyman Rodman, "Marital Power and the Theory of Resources in a Cultural Context", Journal of Comparative Family Studies, Vol. 3 (Spring 1972) pp. 50-67.
- Constantina Safilios-Rothschild, "The Study of Family Power Structure", Journal of Marriage and the Family, Vol. 32, no. 4 (November 1970).
- Constantina Safilios-Rothschild, "A Comparison of Power Structure and Marital Satisfaction in Urban Greek and French Families", Journal of Marriage and the Family, Vol. 29 (May 1967) pp. 345-352.
- William Silverman and Reuben Hill, "Task Allocation in Marriage in the United States and Belgium", Journal of Marriage and the Family, Vol. 29, no. 2 (May 1967) pp. 353-359.
- Daniel Starch and Staff, Male vs. Female Influence in the Purchase of Selected Products, A report prepared for Time Magazine 1958.
- Benjamin Whorf, "The Relation of Habitual Thought and Behavior to Language", in L. Spier et al. Language, Culture and Personality (Menasha, Wisconsin: Banta 1941).
- Donald Wolfe, "Power and Authority in the Family", in Donald Cartwright (ed.) Studies in Social Power (Institute for Social Research, University of Michigan, Ann Arbor 1959) pp. 99-117.

UNDERLYING PERCEPTUAL PATTERNS IN HUSBAND AND WIFE  
PURCHASE DECISION INFLUENCE ASSESSMENTS

Alvin C. Burns, Louisiana State University  
David J. Ortinau (student), Louisiana State University

Abstract

Wives' assessments of the husband and wife purchase decision making influence for thirty product decisions representing five large durable purchases, each with six attendant decisions were subjected to analysis of variance and factor analysis in an exploratory study. The results revealed seven identifiable product-, decision type- and product-decision type-perceptual dimensions in the responses. Individual analysis found two distinct groups of women with respect to these patterns. Perceptual and demographic differences between these two groups are reported and the implications of the findings discussed.

Introduction

Research on the husband and wife purchase decision making process has brought into sharp focus the need for multi-dimensional conceptualizations of the sharing of influence in various purchase decisions. Beginning with the breakthroughs made by Davis (1970), it has become increasingly apparent that husband and wife purchase decision making influence varies across product, across decisions within products, across phases of the decision-making process, and across families. See, for example, Burns, (1975); Burns and Granbois, (1977); Davis, (1976); Davis and Rigaux, (1974); Ferber and Nicosia, (1972); Hempel, (1974); or Woodside, (1974). Most studies, however, stop with this determination and do not continue on to investigate the nature of these variations, nor to suggest patterns which might exist beneath the data. From an intuitive standpoint, one is inclined to suspect that certain tendencies or patterns should persist and these, if discovered, might lead to refined conceptualizations, or at least, insights into the ways in which husbands or wives perceive the manner in which they and their spouses make various purchase decisions.

The study reported in this paper builds on this notion and is an investigation into the nature of husband and wife purchase decision making influence with specific concern for the similarities and dissimilarities in wives' perceptions of influence shared for five representative large durable products and six decisions associated with each of these products. In particular, the study was designed to explore the following questions:

1. Are there differences in wives' perceptions of husband and wife purchase decision making responsibility across products and across decisions for these products?
2. If so, are there underlying patterns in the wives' perceptions of purchase influence distributions which demonstrate and clarify the nature of these differences?
3. If such patterns exist, are there individual differences in the form of subgroupings of women distinct from other subgroupings?
4. If the subgroupings exist, are there distinctions between these groups either in the form of influence perceptions and/or independent variables which may be relied upon as meaningful differences, perhaps useful in directing future research or as methods of

segmenting the total population.

The research questions are not couched as hypotheses given the contingency nature of each question which is based on the results of its predecessor. Also, the next section will reveal other exploratory aspects of this research.

Method

The exploratory study reported in this paper is based on the responses to a questionnaire administered to 81 wives chosen in a nonprobabilistic manner to represent a broad cross-section of women in terms of age, working status, socio-economic position, race, and educational attainment. The convenience sample was drawn primarily from a large Southern city and across that state. The questionnaires were administered at subjects' places of work or in their homes. Although the original study plans included the collection of husbands' responses, data collection difficulties prevented their use in this paper. Nor was the target sample size of 100 wives available for analysis at the time of this writing.

Each respondent was instructed to indicate on a five-point scale of how she believed she and her husband would decide each of thirty randomly arranged purchase decisions in accord with the response categories: "husband alone" (=1); "husband more than wife"; "both equally"; "wife more than husband"; or "wife alone" (=5). The women were directed to respond without conferring with their spouses or other individuals. Respondents were requested to indicate as accurately as possible "the way you believe the decision would be made in your family."

The thirty purchase decisions chosen for investigation in this study were comprised of six decisions for each of five products. The decisions of: when to buy, where to buy, how much to spend, brand or model, style, and color were paired with the large durable goods purchases of: family automobile, dinette set, sofa, television, and stereo set. Large durables purchases such as these are customarily used in husband-wife role structure studies. (Davis, 1976) Moreover, decisions types such as the ones used in this study are typically used in this research, (Davis, 1976) as is the five-point measurement scale (Davis, 1976). A final section of the questionnaire requested details on family and personal demographics.

Inspection of the sample's demographics revealed a diverse group ranging in ages from 21 to 69 years, in length of marriage from 1 to 50 years, in education from grammar school to graduate degree, in wife's workweek from 0 to 50 hours, and a broad range of both husbands' and wives' occupations.

Analysis and Results

Preliminary analysis of these wives' perceptions of the distribution of husband wife influence for the various product decisions was performed by computing means and standard deviations for each of the thirty product-decision combinations. Means and standard deviations were also computed for each product across all six decision types (product means) and for each decision type across

all five products (decision means). These values are displayed in Table 1. The range for these values was found to be from 1.89 (where to buy an automobile) to 3.92 (color of a sofa) with a grand mean of 2.92. Thus while the overall average indicated equal sharing of decision making, the individual decisions paired with products ranged from "husband more than wife" to "wife more than husband." None of the means indicated autonomy. Frequency distributions for the responses to each product-decision (not shown here) verified this result.

various products (across decisions for the products) as well as in deciding various decisions (across products for the decision.) The significant interaction reveals that the wives do not think solely on a product differences basis nor strictly on a decision type differences basis. Instead, for this set of decisions and products, they perceive differences which are specific to individual decisions types for certain products.

In order to ascertain the products and decisions which

TABLE 1  
MEANS AND STANDARD DEVIATIONS FOR THE PRODUCTS AND DECISIONS\*

DECISION	PRODUCT					Totals
	Automobile	Dinette	Sofa	Television	Stereo	
When to buy	2.40 (.74)	3.50 (.84)	3.23 (.91)	2.73 (.72)	2.51 (.84)	2.87 (.44)
How much	2.36 (.95)	3.45 (.81)	2.85 (.86)	2.63 (.88)	2.44 (.89)	2.75 (.53)
Brand	2.20 (.82)	3.45 (.98)	3.54 (.94)	2.20 (.79)	2.17 (.92)	2.74 (.55)
Style	2.30 (.84)	3.82 (.88)	3.70 (.93)	3.05 (.90)	2.43 (.99)	3.06 (.47)
Color	3.01 (.85)	3.86 (.92)	3.92 (.83)	3.28 (.96)	3.01 (.95)	3.42 (.59)
Where	1.89 (.69)	3.34 (1.05)	3.41 (.84)	2.50 (.92)	2.23 (.84)	2.67 (.55)
Totals	2.36 (.54)	3.58 (.62)	3.44 (.60)	2.73 (.55)	2.45 (.61)	2.92 (.88)

\*Based on a five-point scale: 1=Husband Alone, to 5=Wife Alone.

In order to test for the presence of significant differences within the mean values, analysis of variance was performed utilizing the five products and six decisions as two sets of treatment levels and testing for an interaction effect. The results of this analysis are summarized in Table 2 which reveals that significant differences existed between products and between decisions. Furthermore, the interaction of products X decisions was found to be significant. All three cases were significant at the .0001 level.

TABLE 2  
ANALYSIS OF VARIANCE RESULTS  
FOR PRODUCTS AND DECISIONS

Source	S.S.	d.f.	F Value	Significance Probability
Products	603.6	4	194.06	.0001
Decisions	162.5	5	41.8	.0001
Products X Decisions	63.9	20	4.1	.0001
Error	1851.5	2381		
Total	2681.6	2410		

These results indicated that at least one compared difference between the mean values was significant at the .0001 level for: (a) product means, (b) decision type means, and (c) product X decision type means. In short, the wives perceived differences in the influence they (or their spouses) have in deciding the purchase of

were perceived as significantly different from one another, Duncan's multiple range test for significance of differences on means was performed, and the results for the five product means and the six decision type means are contained in Table 3. With regard to the products, the wives perceive the dinette and sofa to be appreciably more under their purchase decision making influence, not strictly and exclusively, but products nonetheless where the wife has greater decision making responsibility than the husband, overall. The stereo and automobile purchases, on the other hand, are appreciably more under the responsibility of the husband on the average, while the television set purchase exists as a separate product where the husband is seen as having only slightly more influence than the wife. In the context of those five products, the television purchase decision defines a joint-decision making region with the other product decisions moving progressively away from this situation.

Across the products for which they are determined, the various decision types demonstrated interesting differences. In particular, the color and style decisions are envisioned, on the average, as distinct from each other as well as distinct from all other decisions. Neither one, however, was found to be greatly under the responsibility of the average wife, with the color decision only slightly so and the style decision falling on the "both equally" point. The four remaining decision types - when, how much, brand, and where - exhibit more husband influence than wife influence in the decision, yet all four means show only a slight move away from joint and equal decision making.

TABLE 3

## RESULTS OF DUNCAN'S MULTIPLE RANGE TEST FOR SIGNIFICANCE OF DIFFERENCES OF MEANS\*

- A. Results for Product Means (Means with the same letter are not significantly different at the alpha level of .01)

Grouping	Mean	Product
A	3.58	Dinette
A		
A	3.44	Sofa
B	2.73	Television
C	2.45	Stereo
C		
C	2.36	Automobile

- B. Results for Decisions Means (Means with the same letter are not significantly different at the alpha level of .01)

Grouping	Mean	Decision
A	3.42	Color
B	3.06	Style
C	2.87	When
C		
D C	2.75	How Much
D C		
D C	2.74	Brand
D		
D	2.67	Where

\*Based on a five-point scale: 1=Husband Alone, ..., 5= Wife Alone.

These comparisons of the significant mean differences with the Duncan multiple range test do not show the differences denoted by the significant interaction term which warns of the probable presence of an "averaging effect" which dampens both the product means and decision means differences. Consequently, a third Duncan multiple range test was performed using the product X decision type classifications. The results revealed considerable overlap of nonsignificant differences between means with groupings or clusters of the thirty cases exhibiting significant differences. Presentation of the results is clumsy and not given here. Instead, a different and more interpretable method, namely factor analysis, was applied to the wives' responses in order to investigate the underlying patterns which were suggested by the analysis of variance, significant interaction term result and third Duncan multiple range test results.

Factor analysis is a convenient method of identifying redundancy in responses based on their variability and simultaneously pointing out groupings of variables which constitute independent response sets. Thus, factor analysis does not concern itself with mean differences; rather, it works with the correlations between the variables under study and combines those highly correlated variables into orthogonal axes or perceptual dimensions in this case. Consequently, R-type factor analysis was applied to this set of wives' responses regarding husband and wife purchase decision making influence for the representative thirty products and decisions. Using the decision rule of cutoff for factors with eigenvalues less than 1.0, eight factors were determined, accounting for 72% of the total variance in the wives' responses.

Varimax rotation was applied to the principal components results to enhance interpretability of the results.

Table 4 contains a summary of the factor analysis results with the decisions arranged by factors. In the interest of interpretability, only those factor loadings with a value greater than or equal to (absolute).40 are reported. This reporting method is consistent with Rummel (1970). It should also be noted that some instances of a product-decision variable loading with (absolute) .40 on two factors did occur, and in these cases, the correlations of that variable to other variables loading on the factors were scrutinized and a judgment was made as to which factor the variable should be associated. This practice of judging spurious results is consistent with Harman(1967).

Seven of the eight factors proved to be readily interpretable through varimax rotation and revealed interesting product-, decision type-, and product-decision type-dimensions in these wives' perceptions of spousal purchase influence. In particular, the results substantiate fairly well the family automobile and the stereo set products as a set of basically homogeneously viewed product purchase decisions. Factor V, for example, includes five of the six decisions for the stereo set while Factor IV accounts for four of the six decisions for the family automobile. The mean values for all of these decision types are quite similar, ranging from 1.9 to 2.5, and indicate that the average husband has greater influence than the average wife for these decisions. Neither of these two factors, however, includes the color decisions, and it is apparent that the color decision occupies a distinct niche in these wives' perceptions, for Factor VIII accounts for the color decisions for three of the five products, including the automobile and the stereo set. The mean values for these three decisions range from 3.0 to 3.3, or primarily equal decision making influence, on the average, as compared to the automobile and stereo set factor decision means.

The remaining factors show a distinction in the type of decision contingent on the type of product for which the decision is being made. In particular, the sofa and dinette set, or furniture products, are consistently paired with each another, yet the decision types are separated into three distinct types. Factor I includes the color and style, or the aesthetic and appearance decisions, for the sofa and dinette set while Factor VI accounts for how much and when, or instrumental decisions, for these two pieces of furniture. Finally, the brand decisions for the furniture products constitute Factor VIII. Inspection of the mean values for the decision types included in these three factors reveals that the aesthetic furniture decisions are those most perceived as under the responsibility of the wife (ranging from 3.5 to 3.9), the instrumental furniture decisions are seen as more equally shared (ranging from 2.9 to 3.5), and the brand furniture decisions are slightly more under the responsibility of the wife than equally shared (both at 3.5).

Of the two remaining factors, Factor II is comprised primarily of instrumental decisions for the television set with the husband perceived as having slightly more responsibility for these decisions than the wife. The final factor, Factor VI, contains the style of the television and the when to buy a new sofa decision. Apart from the fact that both are envisioned as equally shared decisions, the common theme of this factor is not readily apparent from this analysis.

The final phase of analysis concerned an investigation of individual differences in the wives' perceptions of husband and wife purchase decision making based on the perceptual dimensions uncovered by the previous analysis. Consequently, each wife's responses were converted to eight factor scores based on the factor analysis results, and these factor scores were subjected to hierarchical cluster analysis as described in Johnson (1967). The

TABLE 4

## FACTOR ANALYSIS RESULTS FOR PRODUCTS AND DECISIONS

PRODUCT	DECISIONS	FACTOR							
		I	II	III	IV	V	VI	VII	VIII
Dinette	Color	.87							
Sofa	Color	.82							
Dinette	Style	.77							
Sofa	Style	.70							
Dinette	When	.46							
TV	How Much		.81						
TV	When		.73						
Auto	How Much		.69						
TV	Where		.67						
Auto	Style			-.80					
Auto	Model			-.77					
Auto	When			-.74					
Auto	Where			-.48					
TV	Brand			-.41					
Stereo	Brand				.79				
Stereo	Where				.78				
Stereo	When				.74				
Stereo	Style				.67				
Stereo	How Much				.67				
Sofa	How Much					-.84			
Sofa	Where					-.78			
Dinette	Where					-.70			
Dinette	How Much					-.55			
TV	Style						-.77		
Sofa	When						-.54		
Auto	Color							-.71	
TV	Color							-.60	
Stereo	Color							-.53	
Sofa	Brand								.82
Dinette	Brand								.73

cluster analysis results showed that several of the wives represented unique, individual dispositions, although a slight majority did express common tendencies, and a much smaller group aligned together but differently from the larger group. Tests of the significance of differences between the two groups (n=43 and n=11, respectively) were performed to pinpoint the salient distinctions between the two groups' factor scores. With the use of t-tests, significant differences at the .05 level or less were found for Factors III, IV, and VIII, representing the automobile decisions, the stereo decisions, and the furniture brand decisions. Comparison of the mean values for both groups with regard to the decisions accounted for in the three factors revealed that the smaller group consistently credited the husband with greater decision making responsibility in the cases of the six stereo decisions, all of which were found to have significant differences at the .05 level or lower. (The fact that all these decisions were significantly different is predictable, given the factor analysis results.) The remaining differences were not significant beyond the .05 level, but there was evident a consistent tendency for the larger group of wives to indicate greater decision making responsibility for themselves in the furniture brand decisions than did the smaller group of wives.

A final comparison between the two subgroupings was made using demographic variables and difference variables (e.g., differences in spouses' education, ages, incomes and workweeks.) In general, the larger group was found

to be slightly older, married longer, and of higher income. Educational differences were not apparent, but differences in spouses' workweeks (19 and 3.5 hours for the larger and smaller groups respectively, sign. level = .10), number of hours the husband worked per week (42 and 31 hours respectively, sign. level = .11) and number of children (2.2 and 1.4 respectively, sign. level = .07) were found.

## Discussion

The study has important limitations which should be brought to the attention of the reader at the onset of discussion. One obvious constraint on the results and their generalizability concerns the choice of products and decision types. A broader range of both sets could have clarified some of the patterns which emerged from the analysis, although the underlying response tendencies do appear to be fairly definite from most of the analysis here. Another limitation rests in the choice of the sample which is specific to a geographic region and nonprobabilistic in design. The intent was to draw a diverse group of families with the hope that commalities in their perceptions might transcend demographic and other differences. Given this fact, the results are in actuality more profound than one might initially believe. In addition to this consideration, the tone of the paper points out that the use of self-report data such as this pertains to the perceptions of these wives,

rather than to objective observations of the purchase decision making process. Finally, the results are constrained by the assumptions and imperfections of the statistical analysis incorporated herein.

These limitations notwithstanding for this exploratory study, the results of the study offer encouragement and point to the probable nature of wives' perceptions of husband and wife purchase decision making. The authors are of the opinion that this paper has been more than the haphazard application of multivariate statistical techniques to a handy set of data. Instead, the analysis progressed logically, guided by a desire to investigate the patterns of responses and to thus explore the foundations of wives' perceptions of the sharing of purchase decision making influence for a representative set of large durable products and attendant decisions made for these products.

The results vividly demonstrate that an understandable logic does exist beneath the wives' assessments of purchase decision making responsibility for the thirty decisions. As indicated by the significant interaction term in the analysis of variance results, the perceived commonalities slice across some products and across some decisions while at the same time, they are specific to certain products and specific to certain decision types. Subsequent analysis reveals that the stereo set decisions, in particular, appear to be seen as an almost homogeneous set of decisions as are many of the family automobile decisions. Both sets of decisions represent more perceived husband influence than wife influence in comparison to the other decisions in the study. In contrast, the color decision stands as an independent realm of decision making responsibility, separate from other television, stereo and family automobile decisions and more toward joint or equal decision making in these wives' opinions. The final sets of decisions reveal that the two household furniture products, the dinette set and the sofa, are perceived in an identical manner; however, the various decisions for these two products are perceived distinctly. Thus, the aesthetic or appearance aspects of the furniture products constitute a set of decisions wherein these wives believe they have the greatest decision making responsibility, while the instrumental or pragmatic decisions of where and how much to spend are seen as shared more equally. Finally, the brand decisions for the furniture products constitute another dimension of decision making responsibility. These results lend strong support to the conclusion that multidimensional typologies of product types, decision types, and product-decision combinations should be formulated and applied to this area of consumer behavior research. The promise of meaningful results is plainly evident.

The results of the cluster analysis revealed that the wives were not identical in their assessments of the sharing of decision responsibility along the various decision dimensions uncovered by the factor analysis. Interestingly, the two groups identified by the cluster analysis differed on the stereo, family automobile decisions and furniture brand decisions, but not on the other furniture, color, or television-related decisions. Thus, the bulk of the differences in this sample of wives' perceptions of the distribution of husband and wife decision making influence rests along product-specific lines. Given the limitations of this study, these results are insufficiently generalizable to more than hint the possible form of individual differences between the two groups identified by cluster analysis. Still, the fact that such differences were evident with two relatively small groups is encouraging, inasmuch as it represents a conservative test of individual differences in wives' perceptions of purchase decision making responsibility. At this point, it is appropriate to encourage the pursuit of individual differences in wives' perceptions with this

combination of techniques and with the objective of mapping the predispositions of groups on perceptual dimensions.

At the very least, the study has pointed out not only that differences exist but also the dimensionalities of these wives' perceptions of purchase decision making responsibilities shared with their husbands. An interesting observation at this point is in the family automobile product decisions which are envisioned as a very similar group with the exception of the color decision. Historically, the automobile purchase has been a prime target of husband and wife purchase decision making research (See for example, Brown (1961), Burns and Granbois (1977), or Davis (1970). While important findings have occurred in these studies, the results from this study suggest that more fruitful research would eventuate from the inclusion of or concentration on furniture products which represent more sharply differentiated decisions with regard to spousal influence distributions. Similarly, the results indicate that explicit concern for the color and brand decisions for certain products is justified. As one final word of caution, it should be pointed out that factor analysis is an imperfect technique, and wholesale reliance on these results is unwarranted. Certainly, however, the findings in this study suggest some important directions for research in the area of husband and wife purchase decision making.

#### References

- George H. Brown, "The Automobile Buying Decision Within The Family," in Nelson N. Foote, Ed., Household Decision Making. (New York: New York University Press, 1961), 193-199.
- Alvin C. Burns, "Spousal Involvement and Empathy in Jointly-Resolved and Authoritatively-Resolved Purchase Subdivisions," Advances in Consumer Research, Vol. III. Cincinnati, Ohio: ACR, 1975.
- Alvin C. Burns and Donald H. Granbois. "Factors Moderating the Resolution of Preference Conflict in Family Automobile Purchasing," Journal of Marketing Research, 14 (February, 1977), 77-86.
- H. L. Davis, "Dimensions of Marital Roles in Consumer Decision-Making," Journal of Marketing Research, 7 (May, 1970), 168-177.
- H. L. Davis, "Decision Making Within the Household," Journal of Consumer Research, 2 (March, 1976), 241-260.
- H. L. Davis and B. P. Rigaux, "Perception of Marital Roles in Decision Processes," Journal of Consumer Research, 1 (June, 1974), 43-50.
- R. Ferber and F. Nicosia, "Newly Married Couples and Their Asset Accumulation Decisions," in B. Strumpel, J. N. Morgan, and E. Zahn, Eds., Human Behavior in Economic Affairs: Essays in Honor of George Katona. San Francisco: 1972, 161-187.
- H. Harman, Modern Factor Analysis, 2nd Ed. Chicago: The University of Chicago Press, 1967.
- D. J. Hempel, "Family Buying Decisions: A Cross-Cultural Perspective," Journal of Marketing Research, 11 (August, 1974), 295-302.
- Stephen C. Johnson, "Hierarchical Clustering Schemas," Psychometrika, 32 (1967), 241-254.
- R. J. Rummel, Applied Factor Analysis. Evanston, Illinois: Northwestern University, 1970.
- A. G. Woodside, "Dominance and Conflict in Family Purchasing Decisions," in Venkatesan, Ed., Proceedings, Third ACR Conference. Iowa City: ACR, 1972, 650-659



WIVES, HUSBANDS, AND CHILDREN:  
THREE STUDIES OF CONSUMER ROLES

Donald Granbois, Indiana University

Abstract

Despite several aspects of their conceptual frameworks and methodology which could be strengthened, the three papers on consumer roles discussed here offer some insights furthering our understanding of marital roles in decisions and adolescent consumer socialization. Each study suggests a promising direction for further research.

Introduction

In his excellent synthesis and critique of household decision making research, Harry Davis commented that the many studies showing that two or more family members often are involved in joint purchasing and consumption behavior have had little impact on the "mainstream" of consumer research. He argued that this may help explain why associations found between dependent variables reflecting group behavior (such as brand purchase sequence patterns, automobile brands purchased, etc.) and predictor variables (such as attitude structure, personality, and information processing rules) gathered from individuals are often disappointingly weak. (Davis, 1976). One criterion for judging the preceding papers by Burns and Ortinau and by Douglas is suggested by Davis' observation. Put simply, we can ask: "Do the findings and/or hypotheses suggested by the studies imply ways to strengthen research on topics typically structured by models of individual behavior?"

The potential contribution of the paper by Moschis and Moore seems to illustrate Davis' suggestion that we study the consequences of family consumer roles. Their independent variable, household communication structure, is analogous to the role structure types serving as dependent variables in many family studies. The "consequence" (which is seldom pursued in typical consumer family role structure studies) is the impact on the consumer socialization process of children.

One function of the discussion which follows is to react to what seems to be strengths and weaknesses in each paper in terms of their conceptual frameworks, methodologies, and findings. A second, and perhaps more important role of the paper is to elaborate and speculate upon suggestions for further work which are stimulated by the papers.

Husband-Wife Consumer Roles in Two Cultures

Writers in consumer behavior have been far more prone to argue the potential importance of cross-national research studies than to actually conduct them. It is not surprising (but surely disappointing) to find only one truly comparative study in Douglas' literature review. Not only is it of significance from an applied perspective to understand similarities and differences where marketing or other influence programs may be transferred to other cultures by American firms and organizations. It also can be argued that better understanding of patterns in our own culture (especially those so universal as to be "taken for granted") is facilitated by comparative studies highlighting contrasts with distinctly different behaviors. The Douglas study is an ambitious exploratory study on a highly relevant and promising (but sadly under-researched) topic.

Her decisions to work with language-defined groups, to match English-speaking with French-speaking groups (whose behavior might intuitively be expected to be different but not too different), and to use more than one group in each language category all seem commendable and nicely defended in her discussion. Her objective might be described as explanatory: to discover patterns of differences in husbands' and wives' consumer roles which appear to be attributable to cultural differences between the two language groups. The notion of hierarchical objectives shows that explanation is a necessary achievement before prediction and influence can be successfully undertaken, and also warns that description must first be accomplished. Douglas' study makes some contributions toward this necessary description, but does not go far enough to permit more than tentative hypothesizing about French-English cultural differences as explanations for differing family role structures.

This conclusion is clarified by some aspects of her conceptual framework, variable formulation, and research methodology which may have contributed to her somewhat disappointing results. (Data analysis is not an element of any of the three studies requiring much comment. Like many current consumer studies, the papers here illustrate the application of analytical procedures far more sophisticated than the conceptual models and data involved. This state of affairs exists perhaps because consumer researchers benefit from methodological developments borrowed from a large number of behavioral and even non-behavioral disciplines). To discover needed modifications is indeed a major role of exploratory studies, and my comments mostly reflect and occasionally elaborate many things Douglas herself learned from the study.

Her lack of prior hypothesizing as part of developing a conceptual framework is somewhat surprising, given the familiarity of the two cultures of interest. Broad differences in values and resulting consumer behaviors among four English-speaking groups have been summarized by Sommers and Kernan (1967), and one would expect to find similar generalizations about French-language groups in the literature of cultural anthropology. Exploratory research need not be totally atheoretical, and some notion of possible expected differences could have facilitated some desirable narrowing of the scope of the study. The very ambitious framework resulted inevitably in much variability not attributable to cultural differences.

Although it isn't explicitly presented, Douglas' conceptual framework assumes variability in husband-wife roles along several dimensions:

- a) roles may change over time, representing trends such as wives gaining importance due to growing career and work roles;
- b) husbands' and wives' responses may differ due to biases and perceptual differences;
- c) roles may vary across decisions and activities;
- d) role differences may exist among demographic and socio-economic groupings;
- e) role patterns may reflect cultural differences between French- and English-language groups;
- f) role patterns may reflect differences among sub-cultural groups within each language group.

Since the study is the first of its kind, trend changes cannot be addressed. Differences in responses between spouses have been documented and discussed in several American studies. Douglas anticipates this in her design, but does not analyze husband-wife differences. Cultural differences are, of course, the main concern of the study. Commendably, aggregate role differences are not her target; rather, similarities and differences between cultures in each specific activity or decision were sought. Several writers have commented on the weakness of typologies of family role structure based on arbitrarily-selected activities and decisions lumped together to support numerical "indexes" of role structure. However, there is no evidence of systematic thought behind the selection of the 17 behaviors selected. Comprehensive frameworks of consumer behavior activities have been offered by Gredal (1966), Arndt (1976), and Reynolds and Wells (1977), which could perhaps be consulted in revising the list of behaviors asked about.

American research offers support for the notion that role structure differences relate statistically to certain socio-economic and demographic variables (such as family life cycle stage), although nothing like strong, easily-interpreted patterns seem to exist. Here the scope of the study could have been fruitfully narrowed, perhaps by purposely selecting subjects representing a relatively homogeneous segment defined both socio-economically and demographically.

Douglas' key dependent variable is captured by the question, "who was typically responsible," asked for each decision and activity. She follows Davis and others in referring to this as "involvement," a source of possible confusion since other researchers follow Krugman's (1965) conceptualization of involvement as a measure of importance or strength of concern. Besides response categories indicating "mainly husband" and "mainly wife," she (commendably) adapts from typologies of role structure the notion that each individual activity can be conducted jointly or by husband and wife acting individually with about equal frequency. It is unfortunate that pretesting didn't allow the perfection of an unambiguous question format here, since individual decisions conceivably could be performed autonomically for any of several reasons. Explanations may be situational (dry cleaning may be picked up by the person whose schedule for the day provides easiest access to the dry cleaner), may reflect fairness or equity (highly pleasant or strongly disliked activities may be done an equal number of times by each person), or may represent conscious avoidance of conflict that experience may reveal tends to accompany joint decisions or activity performance.

Methodological features weakening the study (which are understandable, given the ambitious scope and exploratory nature of the study) are the small, convenience samples; the very low response rates; and the pooling of husbands' and wives' responses. The latter seems to have little logical justification and may represent a violation of the "independent response" requirement of many statistical techniques.

#### Wives' Role Perceptions in Durables Purchasing

The potential contribution of the research questions pursued by Burns and Ortinau is less apparent than the cross-cultural insights sought by Douglas, and their study is obviously less ambitious in scope. My interpretation of their objective is that they hoped to find explanation through interpreting and attributing meaning to groupings found to exist within essentially descriptive data. Unlike cross-cultural research, where one at least implicitly assumes the values and lifestyles which distinguish one cultural grouping from another will impact on the behavior being studied, the Burns-Ortinau approach starts with no tentative, expected patterns at all. The

researchers indicate in their literature review that variations in role structure related to specific products and sub-decisions within each total purchase decision have been found earlier, but they did not derive hypotheses from the findings of earlier studies for testing. What they have done is to replicate earlier research in a new setting without showing either in their discussion of their research design or in their presentation of results how earlier knowledge is extended, modified, refuted or reinforced by their findings.

However, their paper represents a real contribution by developing a new way to explore patterns of marital participation in sub-decisions associated with durable goods purchases. Since the study is clearly intended to be exploratory, my comments on it should be interpreted as positive suggestions for designing further research, which I hope the authors will undertake.

The conceptual framework here, like that of Douglas, encompasses several decision areas, but is limited to five durable products. The six sub-decisions relevant for each of the products can be classified into the schemes suggested by Gredal (1966) and Arndt (1976). Arndt, for example, distinguishes between budget allocation, generic and variant selection decisions. Gredal's scheme identifies major and minor, individual and collective classifications of durable goods, two distinctions which, if recognized, might have encouraged the researchers to modify their selection of products so that each category would be represented. The study then would have provided the setting for putting to empirical test the role structure hypotheses advanced by Gredal.

The four household goods used include two furniture and two entertainment items, but no appliances. Nelson Foote (1966) has argued the importance of distinguishing between "time-using" and "time-saving," suggesting another desirable modification of products asked about. Except for dinettes, the products chosen all frequently exhibit multiple-unit ownerships within households, a factor likely to complicate the respondents' choice of response. Whose car, whether the television set in question is the family room console or a bedroom portable, whether the stereo is for parents or children, all may have important influence. Similarly, whether the hypothetical purchase would represent a first-time acquisition or a replacement could well affect expected marital roles to the extent these are influenced and modified by experience. Finally, while Douglas' selection of activities seems intuitively to be reasonably concentrated among those where adult preferences could be expected to dominate, three of the five products Burns and Ortinau selected -- auto, television, and stereo -- might well be subject to strong influence by children. While both studies repeat the unfortunate precedent established by earlier studies of ignoring children's roles, the omission seems especially dangerous here.

The dependent variable, wife's perception of "how she believes she and her husband would decide," requires respondents to imagine a hypothetical future event, a practice usually discouraged in discussions of questionnaire design. More troublesome is the ambiguity of the verb "decide." Seventeen years ago, James Morgan's (1961) review of existing role structure studies found nine alternative dependent variables. He concluded that none was satisfactory, since dimensions of power or dominance, motives, process and learning were likely to be confounded by single measures. The variable here surely suffers from this confounding, and one can only speculate about how respondents interpreted the meaning of "decide."

Each response explicitly relates to a single future purchase, complicating the use of a response category indicating an autonomic pattern over all the replications of the decision. However, the very infrequent repetition

of the decisions of interest here makes this a less-serious problem than in the Douglas study.

Finally, the small convenience sample and the extreme heterogeneity of responses are methodological features the researchers themselves emphasize, although one could well argue the study is weakened rather than strengthened by the wide range of ages, education levels, lengths of marriage, occupations, etc. represented by the respondents. The study is far too small to permit rigorous examination of demographic differences.

#### Family Communications and Consumer Socialization

Unlike the others, Moschis and Moore seek to explain individual consumer behavior, but their paper shares the others' concern with family variables. Family communication structure, a typology previously applied in research on the political socialization process, is investigated as a determinant of adolescent consumer socialization and media exposure patterns.

At least two important contributions are promised by research on this topic. Obviously, their findings provide additional understanding of the consumer socialization process of children, showing the role of family influence instead of the more commonly-investigated impact of mass media and peers. To the extent that family communication structures can help explain differences in adolescent socialization (and their paper is suggestive but incomplete in developing plausible reasons why competence is highest in pluralistic homes) this kind of investigation may eventually help in the development of consumer education programs directed at parents as well as children. Better understanding of children's responses to mass media (especially advertising) is of obvious relevance both for advertising management and for regulatory policy.

A second implication relates at least indirectly to the two papers just discussed. The role of children in family consumer behavior is still very poorly understood and often ignored. Moschis and Moore's conceptualization could be extended to predict the role of children's contributions to family decisions in each of the four family types identified in the communications structure paradigm. Such an extension could obviously be facilitated by further understanding of parents' communication and power structures in each of the types. For example, are egalitarian norms and joint marital decision structures most common in pluralistic families, and do adolescents in such homes participate more in consumer decisions for family-oriented products than do children in families of the other types?

The support found for the hypothesized patterns of association between pluralistic structure and four indicators of adolescents' "sophistication" as consumers is provocative. One can't help but wonder what other aspects of adolescent consumer socialization might be similarly related to family communications structure. As in the other studies, the dependent variables here are incompletely explained and appear to be somewhat arbitrarily chosen. Far more investigation will be required before the full implications of pluralistic communications structure are realized.

Further research exploiting the other three cells of the paradigm also appears to be called for. Moschis and Moore provide evidence supporting three hypotheses about the relationships between socio-orientation and motives for media exposure and materialism, which appear to offer useful insights. McLeod and Chaffee (1972) report children from consensual families may be more "open" to persuasive communication, for example, and this is suggestive of only one of many possible directions for future hypothesis development and testing.

#### Further Comments on Needed Research

Douglas has already identified several needed studies in further investigating cross-cultural differences in marital role structure, each of which seems worthy of attention. My paraphrasing of her excellent discussion of these topics results in the following list, ranked in order of priority with respect to desirable timing (not necessarily importance):

- 1) Depth analysis of role norms and underlying beliefs and values found in each cultural group of interest;
- 2) Development of equivalent and comparable categories of socio-economic and demographic categories across the cultural groups to be studied;
- 3) Investigation of hypotheses developed in (1) above through study of carefully-chosen samples drawn from socio-economic and demographic strata in each of the cultural groups to be studied.

Although Burns and Ortinau have not spelled out their ideas for future studies quite as explicitly, the several suggestions made for improving their conceptual framework serve as a program for further developing their study of major decisions. Similarly, the usefulness of Moschis and Moore's FCS paradigm should be explored along the lines suggested in the discussion of their paper.

In all three studies, further conceptual elaboration of the dependent variables was shown to be desirable, and it can only be further emphasized here that concepts and findings from studies of individual consumer behavior may be of great value in this process. Work on information processing, types of evaluation processes, attitude formation and structure, and the relationships between values and choice criteria all may provide help.

The problems of sorting out power, influence, and participation and bias and perceptual problems growing out of self-reporting unfortunately still plague the use of questionnaires in role structure studies. However, it seems we can at least pose our questions so as to deal with explicitly-defined past purchases and intended actual future acquisitions or decisions rather than hypothetical situations. The area still lacks good methodological studies incorporating contrasting techniques. For example, time-activity analysis and observation studies of interaction both appear to be viable ways of studying marital roles in problem-solving and decision-making.

Until means are discovered for providing adequate financial support for research in the area (and here ACR can perhaps someday be involved in a clearinghouse role, bringing together potential contributors to research on specific topics of interest and researchers), we need to be especially careful to avoid overly-ambitious designs. Except when socio-economic or demographic characteristics are the special focus of a study, it makes sense to study one stratum of the population at a time.

#### References

- Johan Arndt, "Reflections on Research in Consumer Behavior," in Beverlee B. Anderson (ed.), Advances in Consumer Research, Vol. III. Cincinnati: Association for Consumer Research, 1976, 213-221.
- Harry L. Davis, "Decision Making Within the Household," Journal of Consumer Research, 2 (March, 1976), 241-260.
- Nelson N. Foote, "The Time Dimension and Consumer Behavior," in Joseph W. Newman (ed.), On Knowing the Consumer. New York: Wiley, 1966, 38-46.

Karen Gredal, "Purchasing Behavior in Households," in Max Kjaer-Hansen (ed.), Readings in Danish Theory of Marketing. Kobenhavn: Einar Harcks Forlag, 1966, 84-100.

Herbert E. Krugman, "The Impact of Television Advertising: Learning Without Involvement," Public Opinion Quarterly, 29 (Fall, 1965), 349-356.

Jack M. McLeod and Steven H. Chaffee, "The Construction of Social Reality," in J. T. Tedeschi (ed.), The Social Influence Process. Chicago: Aldine-Atherton, 1972, 50-99.

James N. Morgan, "Household Decision-Making," in Nelson N. Foote (ed.), Household Decision-Making. New York: New York University Press, 1961, 81-102.

Fred D. Reynolds and William D. Wells, Consumer Behavior. New York: McGraw-Hill, 1977.

Montrose Sommers and Jerome Kernan, "Why Products Flourish Here, Fizzle There," Columbia Journal of World Business, March-April 1967.

## EXTENDING PRODUCT LIFE: TECHNOLOGY ISN'T THE ONLY ISSUE

Margaret DeBell (student), University of Maryland  
Rachel Dardis, University of Maryland

### Abstract

In recent years there has been increased interest in improving the reliability, durability and efficiency of major household appliances. However, information concerning consumer response to such improvements is limited. The purpose of this study was to investigate factors influencing consumer purchase and disposal decisions for major appliances. Two different categories of appliances - washing machines and refrigerators - were selected to examine the impact of fashion/technological obsolescence and performance obsolescence on the disposal decision. The results indicated that factors influencing the purchase or disposal decision varied by product category, and that such variations should be considered in making product durability decisions.

### Introduction

In recent decades a great deal of attention has been directed toward conservation of our natural resources at every level, including household appliance consumption. Figures show that in 1972 an estimated 330 million major appliances were in use in the U.S. A 1973 study by the Center for Policy Alternatives at M.I.T. concluded that both the recovery of resources from appliance discards and the extension of product life had potential value, and recommended further study of both these areas (M.I.T., 1976).

The need for further research is highlighted by the fact that appliances are durable products characterized by high manufacturing costs, high consumer costs both initially and throughout the product life, and long life expectancy. Technological improvements which would increase the life expectancy and energy efficiency of appliances should result in the following benefits: a) a decrease in the demand for scarce resources; b) a decrease in the rate of solid waste disposal; c) a decrease in costs to consumers when initial cost is averaged over the years of product use; and d) a saving to the total economy through the retention of the "value added" during manufacture for products which would otherwise be reduced to scrap or landfill (Lund, 1975). Technologically we know it is possible to manufacture products of increased durability. But whether we can extend product life is another matter. We have witnessed over the past two decades a relatively static discard age. For example, a comparison of service-life expectancies for major household appliances in 1957 and 1975 by Ruffin and Tippet (1975) found the expectancies relatively unchanged. In light of obvious technological advances in this period, such a finding appears surprising. In addition, it appears that technological changes over the past two decades may also have improved product reliability, or the rate of service incidence (M.I.T., 1976). Other factors must therefore be counteracting the impact of technological change on potential product durability.

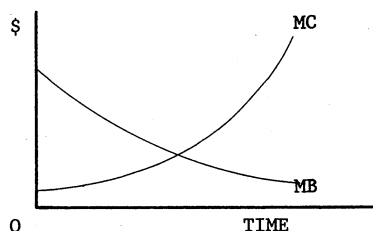
According to Robert Lund of M.I.T., there are several reasons for the unchanged life expectancies of appliances (Lund, 1975). They include: a) high service costs; b) unavailability of replacement parts; c) a mobile population resulting in a high number of second owners of appliances, and; d) consumer affluence

permitting consumers to discard prior to product failure. Technological change resulting in premature product obsolescence or changes in family composition are major reasons for product discard by affluent consumers.

The above discussion is illustrated in Figure 1. The marginal costs and marginal benefits from owning an appliance are measured on the vertical axis, while time is measured on the horizontal axis. When the consumer first purchases the product, the marginal benefits from owning it are high while the marginal costs i.e. operation, maintenance, and repair, are low. Over time the marginal costs are expected to increase, while the marginal benefits decline due to technological or fashion obsolescence. Thus the introduction of automatic defrost refrigerators rendered earlier models obsolete even though performance failure had not occurred. An example of fashion obsolescence is the remodeling of a kitchen and the consequent replacement of working refrigerators due to aesthetic considerations. While the curves are hypothetical and will vary for different appliances and for different consumers, the disposal decision will be made when the marginal costs from ownership exceed the marginal benefits. In terms of the previous discussion, therefore, technological improvements, by extending product life, should shift the marginal cost curve to the right. However, rising repair costs and unavailability of replacement parts due to technological change will shift the marginal cost curve to the left. In addition, technological change will shift the marginal benefit curve to the left by rendering existing products obsolete. As a result, the expected life of appliances may remain relatively static in spite of major changes in product technology.

FIGURE 1

### FACTORS INFLUENCING PRODUCT RETENTION OR DISPOSAL



The objectives of this study were to investigate factors influencing consumer purchase and disposal decisions for major appliances. It was hypothesized that major appliances might be separated into two distinct categories: a) those appliances where mechanical or performance obsolescence was likely to dominate the disposal decision and b) those appliances where fashion or technological obsolescence was likely to dominate the disposal decision. Washing machines, clothes dryers and hot water heaters are likely to fall into the first category, while appliances such as refrigerators, stoves and dishwashers are likely to fall into the second category. Due to their widespread use in American households, washing machines were selected to represent the first category. Both appliances also entail considerable

life-cycle operating costs due to energy consumption, so that respondents' awareness or interest in energy efficiency could also be investigated.

## Results

### Method

#### Sample Section

An important consideration in selecting the sample was to contact persons who had made a purchase or disposal decision within the past two months. This time period was designed to ensure that respondent's answers were based on a recent experience and were hence less likely to result in inaccurate responses due to recall problems.

The target population included both purchasers and disposers of appliances. According to Ruffin and Tippett (1975) approximately 5 - 10% of households are likely to discard a given appliance during a specified year. Thus, a large sample would be required to yield respondents who had made a disposal decision within the past two months. It was concluded that retail stores would be an efficient source of respondents, since the high saturation rate for both refrigerators (99%) and washing machines (72%) meant that a large percentage of appliance purchases were also replacement purchases.

The sample of purchasers was obtained with the help of retailers in the Metropolitan Washington D.C. area. Three types of retailers were selected a) a major discounter, b) a major department store, and c) a specialty appliance store. Each retailer provided us with a list of recent purchasers of washing machines or refrigerators and all persons on this list were contacted.

A telephone survey was used to contact the respondents. However, in order to ensure a favorable response, a letter of introduction explaining the purpose of the survey preceded the telephone call. All surveys were conducted by a single interviewer, thus minimizing interviewer bias. A total of 219 persons were contacted and 179 interviews were completed resulting in an 82% response rate.

#### Questionnaire

The questionnaire consisted of three sections: socio-economic characteristics of respondents, purchase decision process and disposal decision process. The purchase decision section included a) factors influencing brand and store choice, b) search activity by respondents, and c) major product features considered by respondents including durability and energy efficiency. The disposal decision section focused on a) product age at time of disposal, b) reasons for product disposal, c) condition at disposal and d) method of disposal.

In most instances questions were open ended in order to permit maximum flexibility of response. It was also felt that respondents would have little difficulty in answering such questions in view of the relatively short period between the purchase/replacement decision and the interview.

#### Analysis

Chi-Square analysis and rank correlation analysis were used to investigate significant differences between variables. A 0.05 level of significance was used in all instances.

Eighty-nine percent of the refrigerator purchasers and 64% of the washing machine purchasers were making replacement purchases. All respondents were questioned concerning product purchase decisions, while replacement purchasers were questioned concerning product disposal decisions as well.

#### Socio-Economic Characteristics of Respondents

The socioeconomic characteristics of respondents are given in Table 1. As the data indicate, the sample was composed of persons in the older age bracket, and both education and income levels were above average. The fact that the sample is not representative of the U.S. population should be borne in mind in evaluating the results of the survey. It should be emphasized, however, that this was a pilot study which was intended to explore the feasibility of a particular data collection method, and to provide a basis for further research concerning consumer purchase/disposal decisions for major appliances.

There was no significant difference between purchasers of washing machines and purchasers of refrigerators. Thus variations in purchase or disposal decisions could be examined with respect to product category, irrespective of socio-economic characteristics of purchasers.

#### Purchase Decision

There was a significant difference between appliances with respect to reasons for brand choice (Table 2). In the case of refrigerators, features were cited by 57% of the respondents followed by price (44%), brand reputation (31%) and size (26%). In contrast, reliable-in-past ranked first for washing machines (39%) followed by features (38%), brand reputation (31%) and recommendations of consumer magazines (29%). The relative importance of performance/maintenance considerations for washing machines was also shown in the number of respondents citing recommendations of friends or servicemen or service/warranty considerations as reasons for brand choice. Thus while both groups of purchasers were influenced by features, brand reputation and price, service and reliability considerations appeared more important for washing machine purchasers, while features more important for refrigerator purchasers.

Price, brand availability, and convenience were cited as major reasons for store patronage in the case of both appliances. Again, however, service considerations were important for washing machines. Thus product service was cited by 21% of washer purchasers compared to 13% of refrigerator purchasers. In addition 8% of washer purchasers selected the store because it had serviced their old appliance.

A search index was developed based upon procedures discussed by Newman and Staelin (1972) and Katona and Mueller (1955). The search activity index depended on a) number of brands considered, b) number of stores shopped, c) number of information sources, and d) time spent on shopping activity (Table 3). Refrigerators were characterized by higher levels of search activity, probably reflecting the importance of product features and price to consumers.

Features important in appliance selection are given in Table 4. Features common to both washing machines and refrigerators are listed first. Both durability and size are important. However, durability is the major factor for washing machines compared with size for refrigerators. Approximately one third of the respondents cited specific features such as water levels and automatic defrost as important in appliance selection.

Energy efficiency was relatively unimportant for both appliances, ranking seventh for refrigerators and eighth for washing machines.

TABLE 1

SOCIO-ECONOMIC CHARACTERISTICS OF RESPONDENTS\*

Characteristic	Appliance Purchasers	
	Washer Purchasers (n=86)	Refrigerator Purchasers (n=93)
A. EDUCATION	%	%
Part H.S. & Less	3.5	6.4
H.S. Grad - Part College	41.9	43.1
College Grad/Post Grad Work	54.6	48.3
Refusal	----	2.2
B. INCOME		
Under \$20,000	24.4	23.7
\$20-30,000	30.3	25.8
Over \$30,000	39.5	45.1
Don't know	1.1	1.1
Refusal	4.7	4.3
C. AGE		
Under 30 Years	12.8	14.0
30-40 Years	36.0	22.6
40-50 Years	16.3	25.7
Over 50 Years	34.9	36.6
Refusal	----	1.1

\* No significant difference between appliance purchasers at the 0.05 level

TABLE 2

REASONS FOR BRAND CHOICE\*

Reason	Washer Purchasers (n=86)		Refrigerator Purchasers (n=93)	
	%	Rank	%	Rank
Reliable in Past Features	39.5	1	18.3	6
Reputation	38.4	2	57.0	1
Consumer Magazine	31.4	3	31.2	3
Price	29.1	4	23.7	5
Friends	27.9	5	44.1	2
Service/Warranty	24.4	6	9.7	7.5
Other	14.0	7	9.7	7.5
Size	12.8	8	6.5	9
Salesman	10.4	9	25.8	4
Serviceman	8.1	10	5.4	10
	5.8	11	1.1	11

\* Significant difference between ranks at the 0.05 level

TABLE 3

SEARCH ACTIVITY\*

Search Activity	Appliance	
	Washer (n=86)	Refrigerator (n=93)
	%	%
Low	25.6	14.0
Medium	52.3	49.5
High	22.1	36.0

\* Significant difference between appliances at the 0.05 level.

TABLE 4

FEATURES IMPORTANT IN APPLIANCE SELECTION

Features	Appliance	
	Washing Machine %	Refrigerator %
GENERAL		
Durability	66.3	37.7
Size	41.9	71.0
Simplicity	11.7	2.2
Appearance	5.8	19.1
Energy Efficiency	5.8	19.4
SPECIFIC: Washing Machines		
Water Levels	37.3	----
Wash Cycles	33.7	----
Water Temperature	22.1	----
Other (Gadgets)	33.7	----
SPECIFIC: Refrigerators		
Automatic Defrost	----	31.2
Interior	----	30.1
Freezer	----	33.4
Other (Gadgets)	----	23.6

The results are in agreement with previous research by M.I.T. (1976) and the National Bureau of Standards (Stiefel, Kim, Hung, 1976). A 1975 M.I.T. study found the American consumer more interested in features, size, appearance, function and price than in energy consumption. The NBS research report which was issued in 1976 stated that consumers appeared to be relatively indifferent to the operating costs of products relative to the initial purchase price. This "myopic" condition was explained as the result of lack of exposure to quantitative information regarding operating and repair costs, as well as the education to make use of such information.

An energy interest index was developed based on the number of respondents who had considered energy conservation features in the decision process (Table 5). While the majority of the respondents claimed that they had considered energy conservation features, less than one third had considered more than one energy feature. It should also be remembered that respondents were likely to respond positively to queries concerning awareness and interest in energy conservation. Thus the fact that 27% of washing machine purchasers and 42% of refrigerator purchasers had considered no energy feature is significant. The results for refrigerators are particularly important since the M.I.T. study indicated that operating costs (i.e. energy consumption costs) account for more than 50% of the life cycle cost of refrigerators.

TABLE 5

INTEREST IN ENERGY CONSERVATION FEATURES\*

Number of Energy Conservation Features Considered	Appliance	
	Washers (n=86)	Refrigerators (n=93)
	%	%
None	26.7	41.9
One	40.7	28.0
More than One	32.6	30.1

\* No significant difference between appliances at the 0.05 level.

## Disposal Decisions

Disposal has been defined differently by different researchers. At what point is a product discarded? Some will contend that as soon as an owner relinquishes possession of an appliance, it has been discarded. There are problems with this definition, for in some areas the refrigerator and stove must be left with the house, if sold. We, therefore, defined disposal as the action taken by the owner when he decides to replace an appliance, regardless of the motive. For example, if the household's primary refrigerator is replaced by a new refrigerator, the old refrigerator is considered as "disposed" even though it may be retained in the basement. Likewise, if a family sells their home with a refrigerator included in the purchase price, the refrigerator is not considered disposed until the new owner makes a decision to replace it. The disposal decision is therefore a disposal/replacement decision.

As previously mentioned, the two appliances selected for the survey were designed to represent two different categories of household appliances: a) those appliances where performance obsolescence, e.g. mechanical failure was likely to dominate the disposal decision and b) those appliances where new features, size or new appearance was likely to dominate the disposal/replacement decision.

There should be, therefore, differences between washing machines and refrigerators in the condition of the product at the time of disposal and the reasons for disposal. The condition of the appliance at time of disposal is given in Table 6, and confirms this hypothesis. Only 9% of the washing machines were in good working order while 42% did not work at all. In contrast, 31% of the refrigerators were in good working order, while only 17% were inoperative.

TABLE 6

### CONDITION OF APPLIANCE AT TIME OF DISPOSAL/REPLACEMENT\*

Condition	Appliance	
	Washer (n=55) %	Refrigerator (n=83) %
Worked	9.1	31.3
Operating Problems	49.1	51.8
Did Not Work	41.8	16.9

\* Significant difference between appliances at the 0.05 level

The results of Table 6 are reinforced by Table 7 which provides information concerning reasons for product disposal. As might be expected, the great majority of washing machines (93%) were replaced due to mechanical problems, while only 54% of refrigerators were replaced for the same reason. Features were unimportant for washing machines (6%) and accounted for 42% of the product disposal/replacement decisions in the case of refrigerators.

It was also hypothesized that method of appliance acquisition might influence the reasons for disposal, in particular for those appliances where features play a major role in the purchase decision. Thus, second owners of refrigerators, i.e. refrigerators purchased with a house, might be inclined to purchase a new product due to feature inadequacy rather than due to product failure. Since Americans are a mobile society (the 1970 census estimated that approximately 15 million households move each year) this could be expected to make an appreciable contribution to "early retirement"

of appliances, and a supportive case for not increasing appliance durability. A Chi-square analysis showed a significant difference between original purchasers and second owners with respect to reasons for disposal of refrigerators. Thus 54% of the second owners disposed of the product due to inadequate features compared with 29% of original purchasers. In the case of washing machines, there was no significant difference between original purchasers and second owners with respect to reasons for disposal.

TABLE 7

### MAJOR REASONS FOR PRODUCT DISPOSAL/REPLACEMENT\*

Reason	Appliance	
	Washer (n=54) %	Refrigerator (n=80) %
Mechanical	92.7	54.2
Features	5.5	42.2
Other	1.8	3.6

\* Significant difference between appliances at the 0.05 level

The above results are in agreement with a study commissioned by the Canadian government which found that mechanical failure of washing machines was one of the major reasons for product disposal (Stevenson and Kellogg, Ltd.). The same study also surmised that refrigerators and stoves, by virtue of being kitchen appliances, were more likely to be replaced by more modern and decorative units than other appliances, and that new features were of more substantial interest in the purchase of kitchen appliances than other appliances.

The average age of the disposed appliance was 12 years for washing machines and 14 years for refrigerators. These figures are not too dissimilar from those cited by the National Bureau of Standards and the Department of Energy which place the expected life of washing machines and refrigerators at 11 and 15 years respectively.

The disposition of the disposed appliance is given in Table 8. Approximately one-fourth of both appliances were discarded as trash. However, 26% of refrigerators were used elsewhere in the home compared with 4% of the washing machines. This undoubtedly reflects variations in operating conditions at the time of disposal, and the willingness of respondents to discard refrigerators before expiration of their useful life since such appliances continue to provide benefits to the homeowner. One wonders, however, if households have considered the additional cost of operating a second refrigerator. According to the Edison Electric Institute, a 17.5 cubic foot automatic defrost refrigerator consumes 2,250 kwh annually, which corresponds to an annual cost of \$85.50 based on 3.8¢ per kwh.

The results of Table 8 also highlight the importance of determining the disposition of used appliances by dealers. Are used products discarded as trash or are they recycled thus increasing resource use? It is also interesting to note that there were no trade-ins for either appliance. The absence of a second-hand market for appliances makes the disposal decision of the household of prime importance in developing resource conservation policies.



TABLE 8  
METHOD OF APPLIANCE DISPOSAL

Method	Appliance			
	Washer (n=55) %	Rank	Refrigerator (n=83) %	Rank
Dealer Removed	49.1	1	19.3	3
Trash	25.5	2	25.3	2
Gave Away	9.1	3.5	14.5	4
Other	9.1	3.5	2.4	7
Used Elsewhere	3.6	5	26.5	1
Sold	1.8	6.5	7.2	5
Donated	1.8	6.5	4.8	6
Trade-In	---		----	

#### Discussion

According to Robert Lund (1975), the greatest benefit from increased product life is the reduction in consumer expenditures for new appliances. The next major benefit components are the reductions in municipal waste and the use of scarce resources. However, there are possible adverse effects. More durable products might require more expensive materials and might result in a decline in production and in employment. More durable products might also slow down the adoption of energy saving technological innovations. The reservations expressed by Robert Lund were also echoed by U.S. and Canadian manufacturers when they were questioned concerning the feasibility of increasing product durability (Stevenson and Kellogg, Ltd.). While there was agreement that product durability could be increased, the benefits were questioned for the following reasons: a) lack of consumer interest in increased product life, b) increased product cost, c) lack of correlation between potential lifetime and actual lifetime due to replacement of products, and d) use of more energy expensive materials thus defeating any gains in energy conservation.

The results of this consumer survey support many of the reservations of the manufacturers and Robert Lund. It appears that products are discarded by consumers due to technological or fashion obsolescence and increased product durability would have little impact on the age of appliances at discard. However, the results of this study also suggest that there are two different categories of appliances - those that consumers will retain until product failure occurs and those that consumers will discard prior to this condition. Thus product durability was mentioned by 66% of the washing machine purchasers compared to 38% of the refrigerator purchasers. Similarly, 93% of the washing machines were discarded due to mechanical problems, compared with 54% of the refrigerators. This suggests that improvements in product durability would have a greater pay-off for appliances such as washing machines which appear less subject to fashion or technological obsolescence.

While the study did not specifically investigate the availability and cost of repair, a significant number of appliances in both instances were discarded due to operating problems or because they no longer worked. As repair and maintenance costs increase, the benefits from improved product durability may be more than counteracted by the rising costs of maintaining the appliance in good working condition.

The results of this study indicate another area where resource conservation might be achieved - the area of energy conservation. We know that improved energy efficiency is technically feasible and that minimum energy efficiency standards for appliances will probably be formulated within the near future. However, minimum standards are only part of the solution to energy consumption. Consumer awareness of energy costs should result in a demand for more energy efficient appliances, thus increasing resource conservation. Consumers in this survey exhibited relatively little information concerning energy saving features, and energy efficiency ranked low for both appliances with respect to product selection. The proposed energy labeling program which is being implemented by the Federal Trade Commission should do much to correct this situation.

Finally, the particular data collection method used in this study appeared satisfactory. Major advantages included efficiency of data collection and recency of information obtained. While some limitations undoubtedly exist with respect to store participation, the advantages may more than outweigh the disadvantages.

#### References

- George Katona and Eva Mueller, "A Study of Purchase Decisions," in Lincoln H. Clark, ed., Consumer Behavior: The Dynamics of Consumer Reaction (New York: New York University Press, 1955), 30-87.
- Robert T. Lund, "Making Products Live Longer," a paper for the Environment Directorate of the Organization of Economic Cooperation and Development (Paris: Reproduced by M.I.T., 1955).
- M.I.T., "Consumer Appliances: The Real Cost," a report issued by the Center for Policy Alternatives, 1976.
- Joseph W. Newman and Richard Staelin, "Prepurchase Information Seeking for New Cars and Major Household Appliances," Journal of Marketing Research, 9(August, 1972), 249-57.
- Marilyn Doss Ruffin and Katherine S. Tippet, "Service-Life Expectancy of Household Appliances: New Estimates from the U.S.D.A.," Home Economics Research Journal, Vol. 3, No. 3(March 1975), 159-70.
- Stevenson and Kellogg, Ltd., "Major Domestic Appliances and Automobile Tires: Environmental and Economic Impacts of Product Durability," a project report prepared for Waste Management Branch, Environment Canada and Office of Energy Conservation, Energy Mines and Resources, (March, 1977).
- S. Wayne Stiefel, Justin Kim and Howard Hung, "Life Cycle Costing: An Assessment of Practicability for Consumer Products," a report prepared for the National Bureau of Standards, Center for Consumer Product Technology, Institute for Applied Technology, (December, 1976).
- Theodore J. Wang, "Service Lives of Major Appliances," a paper presented at the American Council on Consumer Interests 23rd Annual Conference, April 1977, Denver, Colorado.

#### Acknowledgment

The authors wish to express their appreciation of the help offered by Dr. Theodore Wang and the personnel at the Center for Consumer Product Technology at the National Bureau of Standards. We are also grateful to the Computer Science Center at the University of Maryland, for the use of their facilities.

A PATH-ANALYTIC INVESTIGATION OF LIFE SATISFACTION AMONG ELDERLY CONSUMERS<sup>1</sup>

William O. Bearden, University of South Carolina  
 A. William Gustafson, Texas Tech University  
 J. Barry Mason, The University of Alabama

Abstract

A recursive model depicting a theoretical chain between individual health situation, financial situation, alienation, living level satisfaction, consumer satisfaction, and overall satisfaction with life was examined by path analysis. The data used to test the hypothesized relationships were collected by personal interview from a random sample of 110 elderly consumers living in a medium-sized Southeastern community. Health situation, alienation, and satisfaction with level of living were demonstrated as being plausible determinants of overall life satisfaction.

Introduction

Increasingly, consumer and marketing researchers are directing their attention towards the elderly (e.g., Bernhardt and Kinnear, 1975; Gelb, 1977; Koeske and Srivastava, 1977; Phillips and Sternthal, 1977; Waddell, 1975). This research is largely attributable to the recognition of the aged as a market segment with substantial potential. Growing concerns for the general welfare of the aged in related sociology and gerontology research have resulted in an extensive series of studies designed to isolate correlates of elderly life satisfaction (e.g., Bell, 1974; Bull, 1975; Charfield, 1977; Clemente and Sauer, 1976; Cutler, 1973 and 1973; Edwards and Klemmack, 1973; Knapp, 1976; Lohmann, 1977; Martin, 1973; Spreitzer and Snyder, 1974; Wolk and Telleen, 1976). Only recently have efforts been made to examine changes in life satisfaction over time (Palmore and Kivett, 1977) and to develop causal explanations involving the theoretical determinants of life satisfaction (Medley, 1976).

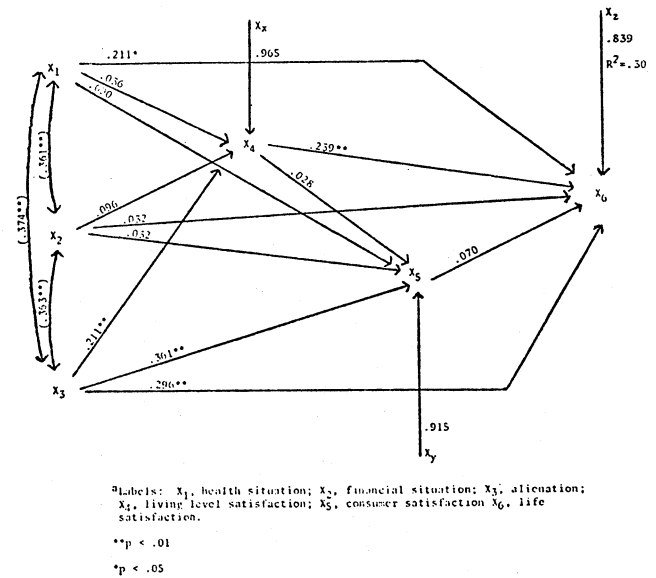
The principal objective of the present study was to investigate a theoretical explanation of elderly consumer life satisfaction. An attempt was made to examine the ability of several frequently hypothesized antecedents of behavior to explain life satisfaction among the elderly. Specifically, a causal chain between health situation, financial situation, alienation, satisfaction with level of living, consumer satisfaction, and overall satisfaction with life was examined via path analysis for a random sample of 110 elderly consumers. These variables have been hypothesized in previous research to impact individual well-being and/or consumer behavior. The intent was to consider directly the fact that many elderly must function as consumers in the marketplace and that a portion of their overall well-being is affected by that ability. Further, an effort was made to provide some additional insight into other consumer-related problems and concerns of the elderly. Specifically, a series of financial problems frequently encountered (Williams, Nall, and Deck, 1976) and the relative priorities of a number of dimensions assumed to underlie quality of life perceptions (Campbell, Converse, and Rodgers, 1976) were examined.

<sup>1</sup>Research funded by a grant from the Administration on Aging, Center for the Study of Aging, The University of Alabama.

Theoretical Determinants of Life Satisfaction

This study does not represent an attempt to develop a complete explanation of elderly life satisfaction but is concerned with examining an extension of the theoretical framework recently suggested by Medley. The modified framework examined is shown in Figure 1.

FIGURE 1



The proposed explanation consists of three exogenous (X<sub>1</sub>, health situation; X<sub>2</sub>, financial situation; and X<sub>3</sub>, consumer alienation) and three endogenous variables (X<sub>4</sub>, satisfaction with level of living; X<sub>5</sub>, consumer satisfaction; and X<sub>6</sub>, overall life satisfaction). Satisfaction with level of living and consumer satisfaction are depicted as intervening measures between the exogenous variables and overall satisfaction with life. The exogenous variables are depicted as being intercorrelated and affecting life satisfaction both directly and indirectly through the two intervening variables (Blalock, 1969; Duncan, 1975). Admittedly, the results would be different with other variables introduced and/or modification of the proposed structure (Medley, 1976: 449). However, previous research does support their inclusion as potential predictors of elderly life satisfaction.

The ability of perceived health conditions and financial adequacy to explain life satisfaction among the elderly has been demonstrated (e.g., Edwards and Klemmack, 1973; Spreitzer and Snyder, 1974; Bull and Aucoin, 1975; Clemente and Sauer, 1976; Palmore and Kivett, 1977). Support has also been found for the hypotheses that higher incomes both influence life satisfaction directly and reduce the impact of health problems on life satisfaction (Chatfield, 1977). Generally, positive feelings regarding household financial adequacy and self-assessed health are assumed to contribute substantially to overall life

satisfaction and successful aging both independently and in interaction with each other.

The concept of alienation has been recently suggested as providing the potential for more fully understanding discontent and dissatisfaction among consumer segments (Pruden and Longman, 1972). Individuals may experience any number of forms of alienation which serve to intensify their discontent and produce frequent criticisms against business (Lambert and Kniffin, 1975). These alternative varieties of alienation (e.g., powerlessness, normlessness, and isolation) are based upon the concepts of values, behavior, and expectations (Seeman, 1961). Alienation is a cognitive state and these forms represent the range of feelings that may result from the interaction of individual values, the impact and nature of individual behavior, and expectations regarding the outcomes of behavior (Fisher, 1976). Further, alienation can be viewed as a relatively enduring experience of dissatisfaction based on these real and expected outcomes (Barakat, 1969) and, as an index of social participation in normal life activities, should predict overall elderly satisfaction with life (Martin, Bengston, and Acock, 1974).

Satisfaction with standard of living has also been shown to significantly influence life satisfaction (Spreitzer and Snyder, 1974). As an example, for elderly persons with high demands on their financial resources, relocation pressures and consumer decisions will likely influence satisfaction with level of living (Nelson and Winter, 1975). Movement from one residence to another is likely to be disorienting and become interrelated with feelings of helplessness and social isolationism (Medley, 1976). However, elderly consumers can influence the satisfaction received from their level of living and from goods and services purchased and consumed. Their ability to shop and operate effectively as consumers is often affected by their health and financial situations and represent viable means of affecting satisfaction with level of living. Satisfaction with level of living should in turn impact both overall life satisfaction directly and indirectly through feelings of consumer satisfaction based on experiences within the marketplace. This may be particularly true for many elderly, since recent reports indicate that those individuals who become dissatisfied but fail to take action and achieve satisfaction tend to be older, more financially pressed, and alienated (e.g., Warland, Herrmann, and Willits, 1975; Andreasen, 1977).

#### Methodology

Data used in this study were collected from 110 elderly households in a medium-sized Southeastern community. Households were selected from a list of all taxpayers filing for homestead exemption within the area. Two hundred households with at least one permanent member 65 or more years of age were randomly selected. Data were collected by personal interview in the respondents' homes by trained interviewers. Each respondent represented the primary grocery shopper within the sampled households. The ninety nonrespondents were distributed among the following categories: refusal (16), absent after repeated callbacks (8), moved (6), deceased (18), both residents under 65 (21), and other (21).

The average age of the primary shopper was 66.2 years. Forty-two percent of the respondents reported household incomes of \$7500 or more and an average education of 9.7 years. The age and education distributions of the present sample are compared in Table 1 with two recently reported national surveys relating to aging:

TABLE 1

Comparison of Age and Education Distributions Across Samples<sup>a</sup>

	1975 CPS <sup>b</sup>	NCOA-Harris <sup>c</sup>	Current
Age			
55-65	--	--	36
65-69	36	37	35
70-74	26	27	18
75-79	18	19	8
80+	20	17	3
Education			
0-7	30	26	24
8	24	21	16
9-11	14	17	19
12	18	19	21
13+	14	17	20

<sup>a</sup>Table entries represent category percentages.

<sup>b</sup>1975 Current Population Survey (Bureau of Census, 1977).

<sup>c</sup>National Commission on Aging - Harris Study (Henretta, Campbell, and Gardocki, 1977).

(1) the 1975 Current Population Survey (Bureau of Census, 1976) and (2) the National Commission of Aging-Harris Study (Henretta, Campbell, and Gardocki, 1977). The percentages shown indicate that the respondents interviewed were somewhat younger and more educated than the elderly described by the national surveys. These differences are most likely attributable to the sampling frame employed (e.g., homeowner couples with at least one spouse 65 or older). While a completely random sample would have included single person households, residents of housing restricted to elderly, and patients in health-care facilities, the selection process did insure a sampling of elderly consumers active within the marketplace.

Twenty-eight percent of the sample indicated that health problems were of concern and limited somewhat their mobility. However, only 11 respondents reported being in poor health. Previous research has shown that physician ratings of individual health and respondent self-evaluations are generally consistent (e.g., Maddox and Douglas, 1973).

#### Operational Measures

Each respondent's current health situation ( $X_1$ ) was operationalized as a composite of their answers to four health-related statements. These statements were designed to assess their self-evaluation of health in terms of their physical mobility, eyesight, hearing, and general health and susceptibility to sickness (Maddox and Douglass, 1973).

The financial situation ( $X_2$ ) of each household was operationally defined as an aggregate score of each respondent's answers to a series of Likert statements reflecting the occurrence of 13 frequently encountered financial problems (see Table 4). The statements were scored four to one and labeled often, sometimes, seldom, and never (Williams, Nall, and Deck, 1976). Specifically, the interviewers used instructions similar to

the following: Aside from not having enough money, which of the following problems do you have and how often do you have this problem?

Feelings of alienation ( $X_3$ ) were assessed by having respondents reply to the 24 5-point Likert "agree-disagree" statements comprising Dean's alienation scale (Dean, 1961; Robinson and Shaver, 1973). A combined score was used to represent the consumer's feelings of alienation.

Elderly satisfaction with their level of living ( $X_4$ ) was operationalized by having each interviewee respond to the following statement (Hafstrom and Dunsing, 1973):

The things people have--housing, car, furniture, recreation, and the way we live--make up their standard of living. How satisfied are you with your present standard of living? That is, with the things you have and the way you are living?

The elderly consumers were asked to respond either very satisfied, somewhat satisfied, somewhat dissatisfied, or very dissatisfied. Responses were scored four to one. This operationalization actually corresponds more closely with level of living since it reflects current levels and is not restricted solely to consumption. "Standard of living", although most often used in individual conversations, generally implies desires and striving for attainment (Davis, 1945).

Consumer satisfaction ( $X_5$ ) was assessed by having each member of the elderly sample indicate their agreement, disagreement, or uncertainty regarding 15 items of the 82 statements comprising the Consumer Discontent Scale recently developed and tested by Lundstrom and Lamont (1976). Eight negative and seven positively worded statements reflecting individual satisfaction with consumer experiences and business practices were selected (e.g., "the quality of goods has consistently improved over the years," "the only person who cares about the consumer is the consumer himself"). Responses to the 15 statements were recorded and combined into an overall score reflecting elderly satisfaction with their interface with the business environment.

Individual overall life satisfaction ( $X_6$ ) was operationally defined by means of the 10-item version of the Life Satisfaction Index developed by Neugarten, Havighurst, and Tobin (1961) and Klemmack, Edwards and Carlson (1974). The Index has been found to provide a reliable estimate of life satisfaction for both elderly and other age groups. The intent of the construct is to capture: zest for life as opposed to apathy; resolution and fortitude as opposed to resignation; congruence between desired and achieved goals; high physical, psychological, and social self-concept; and a happy optimistic mood tone (Adams, 1969). Respondents indicated their agreement, disagreement, or uncertainty regarding seven positively and three negatively worded statements. The instructions used by each interviewer and item statements were similar to the following:

Some people are satisfied with the way they have lived their lives while others may be satisfied with some parts and dissatisfied with others. Please tell me if you agree, disagree, or are uncertain about each of the following statements:

1. The things I do now are as interesting to me as they ever were.
2. I have made plans for things I'll be doing a month or a year from now.

## Results

The simple correlations, means, and standard deviations for each of the model's six components are shown in Table 2. Life satisfaction was significantly correlated as expected with each of the assumed explanatory variables. The potential and actual ranges for each of the variables examined are also shown. The coefficient of determination shown in Figure 1 indicates that the set of antecedent variables accounts for approximately 30 percent of satisfaction in life among the elderly sample ( $F=9.04$ ;  $df=5,104$ ;  $p<.01$ ). The resulting path coefficients and their level of significance are shown on each single-direction arrow while the simple correlations between the three hypothesized exogenous variables are shown in parentheses.

Comparisons of the explanatory variables in terms of their direct effect on life satisfaction indicate that alienation and satisfaction with level of living are significant ( $p<.01$ ) in their relationship with overall life satisfaction. A significant path was also found between alienation and life satisfaction through level of living satisfaction. The effects, if any, of the financial situation of each household on overall life satisfaction were felt through either respondent feelings of alienation or health concerns. Consumer satisfaction was not found to impact directly life satisfaction among the elderly.

A more complete decomposition of the resulting path relationships is presented in Table 3. The total effect between consumer satisfaction and overall life satisfaction as represented by the zero-order correlation is largely the result of a sizable spurious effect. However, health situation and alienation seemed to influence life satisfaction both directly and in conjunction with level of living satisfaction. The effect of the marginal significance ( $p<.05$ ) between each household's financial situation and the respondent's overall satisfaction with life was apparently accounted for by the remaining variables.

A key idea in path analysis is that path coefficients can be used to estimate the empirical correlations among the variables in the system. Algebraic and graphical procedures for expressing each correlation as a function of path coefficients are discussed by Duncan (1966) and Heise (1969).<sup>2</sup> The reproduced correlations shown above the main diagonal in Table 2 are generally consistent with the sample based variable correlations. The similarity between the two sets of correlations is encouraging and indicates that the proposed framework adequately fits the data. Actually, the procedure of reproducing variable correlations via estimated path coefficients and then comparing with actual correlations is more appropriate for rejecting existing theory than developing new empirically based theory. However, the closeness between the actual and the estimated correlation coefficients does support the proposed framework.

<sup>2</sup>The mathematical expressions used to reproduce variable correlations consisted of a single direct path plus the sum of the compound paths representing all the direct connections allowed by the diagram. The following rule was used: Read back from variable  $i$  then forward to variable  $j$  forming the product of all paths along the traverse, then sum these products for all possible traverses (the same variable cannot be intersected more than once in a single traverse; in no case can one trace back having once started forward; bidirectional correlations are used in tracing either forward or back but only one can be used in a single traverse) (Duncan, 1966, p. 6).

TABLE 2

Variable Correlations, Means, Standard Deviations, and Ranges

Variable	X1	X2	X3	X4	X5	X6	Mean	S.D.	Poten- tial Range	Actual Range
X1: Health Situation	1.000	.361	.374	.148	.241	.376	12.35	2.50	4-17	6-17
X2: Financial Situation	.361***	1.000	.363	.185	.199	.274	43.06	8.92	13-52	19-52
X3: Alienation	.374***	.363***	1.000	.256	.414	.477	72.08	12.16	24-120	42-96
X4: Living Level Satisfaction	.150*	.186**	.259***	1.000	.144	.354	2.59	.63	1-4	1-4
X5: Consumer Satisfaction	.211**	.129*	.377***	.074	1.000	.277	31.34	3.80	15-45	24-41
X6: Life Satisfaction	.361***	.205**	.452***	.346***	.240***	1.000	16.06	3.37	0-20	4-20

The correlations below the diagonal were estimated from sample data. The correlations above the diagonal were derived using the estimated parameters of the model.

\*\*\*p .01  
 \*\*p .05  
 \*p .10

TABLE 3

Decomposition of Path Relationships

Theoretical Path Effects	Coefficients
Consumer Satisfaction - Life Satisfaction	
Total	.240 $r_{65}$
Direct	.070 $r_{65}$
Spurious	.170
Level of Living Satisfaction - Life Satisfaction	
Total	.346 $r_{64}$
Direct	.239 $r_{64}$
Spurious	.105
Indirect (By Consumer Satisfaction)	.002 $r_{54} \cdot r_{65}$
Health Situation - Life Satisfaction	
Total	.361 $r_{61}$
Direct	.211 $r_{61}$
Joint	.135
Indirect	.015
(By Consumer and Living Satisfaction)	.001 $r_{61} \cdot r_{54} \cdot r_{65}$
(By Consumer Satisfaction)	.006 $r_{51} \cdot r_{65}$
(By Living Level Satisfaction)	.009 $r_{41} \cdot r_{64}$
Financial Situation - Life Satisfaction	
Total	.205 $r_{62}$
Direct	.032 $r_{62}$
Joint	.147
Indirect	.026
(By Consumer and Living Satisfaction)	.001 $r_{42} \cdot r_{54} \cdot r_{65}$
(By Consumer Satisfaction)	.002 $r_{52} \cdot r_{65}$
(By Living Level Satisfaction)	.023 $r_{42} \cdot r_{64}$
Alienation - Life Satisfaction	
Total	.452 $r_{63}$
Direct	.296 $r_{63}$
Joint	.080
Indirect	.076
(By Consumer and Living Satisfaction)	.001 $r_{43} \cdot r_{54} \cdot r_{65}$
(By Consumer Satisfaction)	.025 $r_{53} \cdot r_{65}$
(By Living Level Satisfaction)	.050 $r_{43} \cdot r_{64}$

TABLE 4

Frequency of Elderly Financial Problems

Financial Problem <sup>a</sup>	Some-				Ranking <sup>b</sup>
	Often	times	Seldom	Never	
Not enough food until payday	7	8	11	84	8
Behind on house or rent payment	--	2	2	106	11
Unable to buy for recreation	22	13	12	63	2-3
Not enough money for medical care	22	12	4	72	4
Danger of having utility service terminated	4	4	2	100	10
Unable to meet large bills	9	9	3	89	9
Cannot maintain household appliances and equipment	6	11	11	82	7
Unable to afford new clothes and shoes	9	14	9	78	6
Not able to save for emergencies	40	11	10	49	1
Someone else spends money first	--	--	--	110	13
Money is lost, stolen, or taken	--	--	1	109	12
Unable to fix or buy new furniture	24	15	8	63	2-3
Not enough money for personal care and grooming	12	14	8	76	5

<sup>a</sup>Source: Williams, Nall, and Deck (1976).

<sup>b</sup>Ranking based on total number of positive responses (e.g., seldom, sometimes, often).

The frequency of medical-related financial demands is likely to be a particular cause for elderly concern.

### Discussion

Although financial situation did not apparently impact elderly life satisfaction directly, an examination of the specific problems frequently faced by elderly households and assumed to represent the financial situation incorporated in this study provides some insight into the practical concerns affecting aged consumers. These problems and their relative importance based on the responses of the present sample are shown in Table 4. The rankings shown reflect the relative frequency of each problem's occurrence. The resulting order of problems indicates the tendency of the elderly to shift their expenditures away from purchases that can be postponed (e.g., recreation, savings, repairs) to avoid financial problems with more pressing daily needs (e.g., rent and utility payments).

This study attempted to examine the ability of several frequently discussed antecedents of behavior to predict and explain overall elderly satisfaction with life. A theoretical framework linking overall life satisfaction as the principal dependent variable to health situation, financial situation, alienation, consumer satisfaction, and level of living satisfaction was examined via correlation and path analysis. An individual's health situation, feelings of alienation, and satisfaction with level of living were found to be related with overall satisfaction with life. However, only 30 percent of the variance was explained by the variables.

Certainly, further research is needed to more adequately identify the causal determinants of overall life

satisfaction before substantive conclusions can be drawn. These results would likely have been different if additional variables were included or the structure of the antecedents of life satisfaction assumed differently. For example, alienation might have been easily positioned as a result of health, financial situation, and age and not an exogenous measure with elderly health and financial situations as depicted here. Further, a larger sample of elderly including those more restricted by health and living environment could have altered the relative influences of the variables examined. Specifically, the use of only home-owner couples who are much more likely to have been active shoppers may account for the insignificance of the consumer satisfaction measure in the results of this study. Lastly, the use of different operational measures designed and validated for the study of individual life satisfaction is needed. Multiple item measures of level of living satisfaction and use of the complete CDS scale, while further lengthening the interviews, would have improved variable measurement.

Related to overall life satisfaction is the notion of quality of life. Campbell, Converse and Rodgers (1976) have recently suggested items that are principal contributors to individual perceptions of quality of life. Priority rankings were assigned to the items by the 110 elderly interviewed in the present study as a follow-up to the life satisfaction effort. The relative rankings somewhat support the findings of this study and indicate other measures as potential determinants of life satisfaction. Specifically, the importance attributed to family life might explain the significance of alienation in this study and suggests the direct consideration of family life satisfaction in future efforts. Health and financial concerns were considered relatively important and should be considered in future life satisfaction/quality of life studies. Of the quality of life items examined, personal consumption aspects (e.g., furniture and household items) were generally not regarded as being critical when compared with the important concerns regarding home life and individual security and well-being.

#### References

- Abdel-Ghany, Mohamed (1977), "Quality of Life from the Perspective of the Elderly," Home Economics Research Journal, 6, 38-47.
- Adams, David L. (1969), "Analysis of a Life Satisfaction Index," Journal of Gerontology, 24, 470-474.
- Andreasen, Alan R. (1977), "Consumer Dissatisfaction and Market Performance." Faculty Working Paper #377, College of Commerce and Business Administration, University of Illinois, Urbana, Champaign.
- Barakat, Halim (1969), "Alienation: A Process of Encounter Between Utopia and Reality," British Journal of Sociology, 20, 1-10.
- Bell, Bill D. (1974), "Cognitive Dissonance and the Life Satisfaction of Older Adults," Journal of Gerontology, 29, 564-571.
- Bernhardt, Kenneth L. and Kinneer, Thomas C. (1975), "Profiling the Senior Citizen Market," in B. B. Anderson (ed.), Advances in Consumer Research, Vol. 3. Cincinnati: Association for Consumer Research, 449-452.
- Blalock, Hubert M. (1969), Theory Construction, Englewood Cliffs, New Jersey: Prentice Hall, Inc.
- Bull, C. Neil and Aucoin, J. B. (1975), "Voluntary Association Participation and Life Satisfaction: A Replication Note," Journal of Gerontology, 30, 73-76.
- Campbell, Angus, Converse, P. E. and Rodgers, W. L. (1976), The Quality of American Life: Perceptions, Evaluations, and Satisfaction, New York: Russell Sage Foundations.
- Chatfield, Walter F. (1977), "Economic and Sociological Factors Influencing Life Satisfaction of the Aged," Journal of Gerontology, 32, 593-599.
- Clemente, Frank C. and Sauer, William J. (1976), "Life Satisfaction in the United States," Social Forces, 54, 621-631.
- Cutler, Stephen J. (1972), "The Availability of Personal Transportation, Residential Location, and Life Satisfaction Among the Elderly," Journal of Gerontology, 27, 383-389.
- \_\_\_\_\_ (1973), "Voluntary Association Participation and Life Satisfaction: A Cautionary Note," Journal of Gerontology, 28, 96-100.
- Davis, Joseph S. (1945), "Standards and Content of Living," American Economic Review, 35, 1-15.
- Dean, Dwight G. (1961), "Alienation: Its Meaning and Measurement," American Sociological Review, 26, 753-758.
- Duncan, Otis D. (1975), Introduction to Structural Equation Models, New York: Academic Press, Inc.
- \_\_\_\_\_ (1966), "Path Analysis: Sociological Examples," American Journal of Sociology, 72, 1-16.
- Edwards, John N. and Klemmack, David L. (1973), "Correlates of Life Satisfaction," Journal of Gerontology, 28, 497-502.
- Fisher, Claude S. (1976), "Alienation: Trying to Bridge the Chasm," British Journal of Gerontology, 27, 35-49.
- Gelb, Betsy D. (1977), "Gray Power: Next Challenge to Business," Business Horizons, 20, 38-45.
- Hafstrom, Jeanne L. and Dunsing, Marilyn M. (1973), "Level of Living: Factors Influencing the Homemaker's Satisfaction," Home Economics Research Journal, 2, 119-132.
- Heise, David R. (1969), "Problems in Path Analysis and Causal Inference," in E. F. Borgatta (ed.), Sociological Methodology. San Francisco: Jossey-Bass, Inc.
- Henretta, John C., Campbell, Richard T. and Gardocki, Gloria (1977), "Survey Research in Aging," The Gerontologist, 17, 160-167.
- Klemmack, David L., Edwards, John N., and Carlson, J. R. (1974), "Measures of Well-Being: An Empirical and Critical Assessment," Journal of Health and Social Behavior, 15, 267-270.
- Koeske, R. D. and Srivastava, R. K. (1977), "Preliminary Examination of the Sources and Handling of Consumer Problems by the Elderly." Faculty Working Paper, Graduate School of Business, University of Pittsburgh.
- Knapp, Martin R. (1976). "Predicting the Dimensions of Life Satisfaction," Journal of Gerontology, 31, 595-604.

- Lambert, Zarrel V. and Kniffin, Fred W. (1975), "Consumer Discontent: A Social Perspective," California Management Review, 18, 36-44.
- Liu, Beu-Chieh (1975), "Quality of Life: Concept, Measure and Results," American Journal of Economics and Sociology, 34, 1-13.
- Lohmann, Nancy (1977), "Correlates of Life Satisfaction, Morale, and Adjustment Measures," Journal of Gerontology, 32, 73-75.
- Lundstrom, William J. and Lamont, Lawrence M. (1976), "The Development of a Scale to Measure Consumer Discontent," Journal of Marketing Research, 13, 373-381.
- Maddox, George L. and Douglass, Elizabeth B. (1973), "Self-Assessment of Health: A Longitudinal Study of Elderly Subjects," Journal of Health and Social Behavior, 14, 87-93.
- Martin, William C. (1973), "Activity and Disengagement: Life Satisfaction of In-Movers into a Retirement Community," The Gerontologist, 13, 224-227.
- \_\_\_\_\_, Bengston, V. L. and Acock, A. C. (1974), "Alienation and Age: A Content-Specific Approach," Social Forces, 53, 266-273.
- Medley, Morris L. (1976), "Satisfaction With Life Among Persons Sixty-Five Years and Older," Journal of Gerontology, 31, 448-455.
- Nelson, Linda M. and Winter, Mary (1975), "Life Disruption, Independence, Satisfaction, and the Consideration of Moving," The Gerontologist, 15, 160-164.
- Neugarten, Bernice L., Havighurst, R. J., and Tobin, S. S. (1961), "The Measurement of Life Satisfaction," Journal of Gerontology, 16, 134-143.
- Palmore, Erdman and Kivett, Vira (1977), "Change in Life Satisfaction: A Longitudinal Study of Persons Aged 46-70," Journal of Gerontology, 32, 311-316.
- Phillips, Lynn W. and Sternthal, Brian (1977), "Age Differences in Information Processing: A Perspective on the Aged Consumer," Journal of Marketing Research, 14, 444-457.
- Pruden, Henry O. and Longman, Douglas S. (1972), "Race, Alienation, and Consumerism," Journal of Marketing, 36, 58-63.
- Robinson, John P. and Shaver, Phillip R. (1973), Measures of Social Psychological Attitudes, Ann Arbor, Michigan: Institute for Social Research, 275-278.
- Seeman, Melvin (1959), "On the Meaning of Alienation," American Sociological Review, 24, 783-791.
- Spreitzer, Elmer and Snyder, E. E. (1974), "Correlates of Life Satisfaction Among the Aged," Journal of Gerontology, 29, 454-458.
- U.S. Bureau of Census (1976), "Demographic Aspects of Aging and the Older Population in the United States: Current Population Reports." Special Studies Series P-23, No. 59, USGPO, Washington, D.C.
- Waddell, Frederick E. (1975), "Consumer Research and Progress for the Elderly--The Forgotten Dimension," Journal of Consumer Affairs, 6, 164-175.
- Warland, R. H., Herrmann, R. O. and Willits, Jane (1975), "Dissatisfied Consumers: Who Gets Upset and Who Takes Action," Journal of Consumer Affairs, 6, 148-163.
- Williams, F. L., Nall, Martha, and Deck, P. Z. (1976), "Financial Problems of Urban Families," Home Economics Research Journal, 4, 185-196.
- Wolk, Stephen and Telleen, Sharon (1976), "Psychological and Social Correlates of Life Satisfaction as a Function of Residential Constraint," Journal of Gerontology, 31, 89-98.
- Wright, Sewell (1960), "Path Coefficients and Regression Coefficients: Alternative or Complimentary Concepts," Biometrics, 16, 189-202.

SMOKING BEHAVIORS AS A DIFFUSION PROCESS WITHIN AGE COHORT GROUPS:  
AN APPLICATION OF THE SOCIETAL MARKETING CONCEPT

Theodore F. Smith, Old Dominion University  
Richard W. Olshavsky, Indiana University  
Michael F. Smith (student), Indiana University

Abstract

This paper characterizes smoking behaviors in the context of an innovation diffusing through sequential groups of age cohorts. The authors conclude that diffusion theory has considerable potential for adding new, potentially effective dimensions to anti-smoking campaigns as well as campaigns for other products and services subject to age-cohort diffusion.

Introduction

For several years, many marketers have been concerned with the relationships between marketing and society, and particularly the effect marketing programs have on the quality of life. The origins of increased concern parallel in many ways the factors underlying the increasing interest in the current consumerism movement. These factors include: (1) the increased tendency to question all institutions as a legacy of the social upheaval of the 1960's, (2) increased concern with environmental affairs, including both pollution and conservation, and (3) depending on one's viewpoint, either the failure of the marketing concept or its effectiveness.

Kotler's (1977) advocacy of a societal marketing concept provides a rationale for involvement in social issues, such as anti-smoking campaigns. The goal of this involvement parallels the marketing concept; i.e., better serving the long term interests of consumers, not just their felt needs. Furthermore, an application of diffusion theory to this goal may be seen as a response to Rogers' criticism that most diffusion research can be typified as pro-innovation (Rogers, 1976). Such a goal would also be consistent with Zaltman's admonition for a greater sociological input into consumer behavior research (Zaltman and Wallendorf, 1977).

If one accepts the proposition that advertising played an indirect if not direct role in establishing the popularity of smoking behaviors, it would seem appropriate that marketers should be involved in the diffusion of the theme underlying anti-smoking campaigns. As Wiebe (1951) and others have pointed out, however, several prerequisites have to exist for mass communications to effect social change. Since these prerequisites are not often present, social marketing is much more complex than merely a shift of focus as suggested by Farmer (1977).

Nevertheless, by applying diffusion theory to the study of the spread of smoking behaviors among groups of age cohorts, at the very least we should be able to glean new insights into the process of the diffusion of smoking behavior. Ultimately, such an extension of diffusion theory may also generate new approaches that will improve the effectiveness of anti-smoking campaigns.

In the following sections, the spread of smoking behaviors is related to the elements of the diffusion process wherein some possible further applications of reverse diffusion are presented.

The Spread of Smoking Behaviors  
As a Diffusion Process

In his formative work, Rogers (1962) segmented the diffusion process into four elements: (1) the innovation, (2) the communication of the innovation from one individual to another, (3) the social system in which innovations spread, and (4) the period of time during which the diffusion occurs. An analysis of smoking behaviors, on the basis of each of these four elements, appears to be a useful basis upon which to explore the potential for diffusion theory applications to the spread of smoking behaviors.

The Innovation

Although some disagreement exists as to what constitutes an innovation, if we recognize the importance of perceptual differences in consumer behavior, it seems logical to conclude that the most useful definition of an innovation would include anything that is seen as new. Thus, cigarette smoking was an innovation when cigarettes were first introduced into the United States from Europe around 1980 (Robert, 1969). In a real sense, moreover, smoking is an innovation for each group of age cohorts as they enter early adolescence and individually decide, albeit often under various types of social pressures, whether or not they are going to smoke.

On the basis of the impact of this innovation on the social system, smoking behaviors can be classed as a discontinuous innovation, both when originally introduced into the United States and when adopted by the first members of an age-cohort group even today. According to Robertson, a discontinuous innovation involves a significant alteration of behavior patterns, including the learning of new behaviors (Robertson, 1967). The new behaviors that must be learned today, as when cigarettes were first introduced, range from learning to smoke without coughing to learning how to avoid fires while smoking in bed.

The adoption of cigarettes can also have disruptive effects on social relationships, which is another effect characterizing discontinuous innovations. Non-smokers may be repulsed by the smoke and thus by the smoker. Consequently, smokers may seek out other smokers for social support of their habit and, at the same time, gradually reduce associations with non-smokers, especially judgmental ones.

The Communication Process

The communication process involves the performance of necessary functions by key change agents which includes the innovators or early adopters and the influentials and opinion leaders. In the public health field, for instance, attention has been devoted to identifying the profiles of children who smoke early, identified as the innovators or early adopters of cigarettes within their



respective age cohort groups (Williams, 1972). Regrettably, less attention has been focused on the exact role of the influentials or opinion leaders in relation to the diffusion of cigarette smoking.

The influentials or opinion leaders can be subdivided into groups characterized as positive and negative, as well as professional and personal. The positive, professional influentials for cigarettes would include celebrities who explicitly endorse specific brands of cigarettes or, who, by their public smoking behaviors, tacitly provide a generic endorsement. For example, the cowboys in the macho ads for Marlboro cigarettes serve as influentials, underscoring the association between masculinity and cigarette smoking.

For the purposes of anti-diffusion of cigarettes, however, it appears that the personal influentials need more attention. Previous research has identified possible motivations for personal influentials in the diffusion of other types of products. Among the suggested motivations for the behavior of personal influentials are the following: (1) product involvement; (2) self-fulfillment; (3) concern for others; (4) message involvement; and (5) dissonance reduction. A more thorough understanding of why positive influentials and opinion leaders encourage smoking within their peer groups may lead to suggestions for ways to reduce the incidence of this behavior or to minimize its effects.

A negative influential or opinion leader would be one who is effective in discouraging adoption of a product or service. In the context of retarding the diffusion of cigarette smoking among youths, a negative influential would be someone whose opinions are respected by the target population and who communicates that it is "not cool" to smoke. That the would-be influential's opinions be respected is an important prerequisite for effective influence, lest the campaign fail due to a behavioral application of a mathematical principle: a minus times a minus equals a positive. That is, a negative message from a source that is disrespected would be perceived as positive, not negative, reinforcement for the behavior in question. This same principle can be turned around, however, by having a disrespected source give a positive (pro-smoking) message to achieve a negative impact. A classic example of this latter case is the picture of the skid row bum with a butt curled over his lower lip, and the caption "smoking is glamorous."

#### The Social System

The social climate or milieu for the diffusion is the third element of the diffusion process. For traditional product or service diffusion studies, the focus has been on such factors as the extent to which the prevailing social mores and folkways favor change in the subject area. Viewing smoking behaviors as an innovation in the process of diffusion, some relatively recent social changes are noteworthy. Until recently, smoking was a dominant behavior, at least in some settings; e.g., cocktail parties, coffee shops, and in business conference rooms. Currently, however, smoking appears to be assuming the aura of a deviant behavior, especially in adult society. The task from an anti-diffusion viewpoint, is how to help establish smoking as a deviant behavior in the eyes of teenagers. Thus, the social system is an important facet of the anti-diffusion of smoking not in the usual sense of the willingness of people in the system to accept new products or patterns of behavior, but rather the social system is important more in terms of the extent to which it incorporates attitudinal underpinnings of smoking behaviors.

#### The Time Required

The time span during which the diffusion occurs is the fourth element of the diffusion process. With conven-

tional diffusion theory applications, the concern here is whether or not the diffusion pattern follows the usual S-shaped curve. The concern in the case of the anti-diffusion of cigarettes, however, is to impede the diffusion of cigarette smoking within each successive group of age cohorts. Such a retardation of the process is considered important, because some empirical evidence suggests that people who begin smoking relatively early in their lives have a much more difficult time quitting the habit. Also, the earlier individuals begin smoking the sooner in their lives the cumulative, deleterious effects of cigarette smoking might lead to serious disorders.

#### Adoption Process

Diffusion researchers also concern themselves with the decision process, at the individual level, which precedes adoption--i.e., the evaluative criteria used, sources of information, extent of search for alternatives, and so on. Considerable research on cigarette smokers (Williams, 1972) particularly with adolescents, suggests that smoking is initiated for one or more reasons: imitation and emulation of adults or peers, curiosity, and simple acquiescence to group pressure. As such, it appears that no real decision process precedes the sampling of the first few cigarettes. Thereafter, the rapid development of physiological dependence appears to make decision making an ad hoc activity.

#### Suggestions for Additional, But Explicit, Applications of Diffusion Theory In Anti-Smoking Campaigns

Thusfar, we have explored the bases upon which the spread of smoking behavior can be viewed as a diffusion of an innovation. In this section, some explicit possibilities for additional diffusion theory applications are presented. To organize the suggestions for additional ways of applying diffusion theory to this anti-diffusion problem, the discussion will once again be subdivided according to Rogers' four elements of the diffusion process.

#### The Innovation

Smoking behavior can be typified as a recurring innovation-adoption process which is diffused through successively new and unaware age cohorts. The cycle depicting the evolution of non-smoking children to heavy user teenagers has served to facilitate the extension of the product life cycle for cigarettes. It is proposed that a reverse innovation-adoption process be instigated wherein non-smoking is the innovation to be diffused through new age-cohorts. Utilizing diffusion theory, the goal of the campaign is to create a rapid decline in the life cycle of cigarettes by focusing on each new wave of prospective smokers. Such a strategy has been termed a countermarketing strategy by Kotler (1977).

As a possible source of potentially effective appeals for an anti-diffusion campaign, Rogers and Shoemaker's paradigm of product characteristics associated with early adoption of a product is worthy of consideration (Rogers and Shoemaker, 1971). These characteristics are applied in reverse as follows:

1. Relative advantage - The anti-diffusion campaign could stress further the advantages of not smoking or of smoking pipes or cigars in lieu of cigarettes.
2. Compatibility - The incompatibility of smoking with existing values such as preservation of one's body or the incompatibility of smoking with other concerns such as independence, romance, and pollution control.

3. Divisibility - The difficulty of smoking on a trial basis due to the dangers of rapid physiological and psychological addiction.

4. Communicability - Health dangers are difficult to convey to adolescents, for most of the serious damage of smoking arise after many years of smoking, and that prospect is too remote to be real to adolescents, especially in view of the common adolescent sense of immortality.

5. Complexity - The implications of smoking on other aspects of life ranging from greater need (expense) for smokers for cleaning services to higher incidence of respiratory infections, with the attendant restrictions on other activities during periodic recuperations. In other words, emphasizing that a decision to smoke impacts many areas of life, due to complex inter-relationships.

#### The Communication Process

Some actions have already been taken utilizing communications media to reduce smoking behaviors. Examples of such moves would include the anti-smoking television commercials, the ban on broadcast cigarette commercials, and the moral suasion of television personalities to reduce, if not eliminate, on-camera smoking. It would appear, however, that the mass media may be efficient in the sense of creating awareness of the existence of a problem, but moving people to action requires interpersonal reinforcement. Lazarsfeld and Merton have labeled this interpersonal function "supplementation" (Lazarsfeld and Merton, 1949). Greater attention must be directed to using the interpersonal communications process to encourage discontinuance of smoking behaviors or, in a positive sense to promote the diffusion of non-smoking as the social norm. Given the difficulty smokers encounter when they try to quit, it is recommended that better ways be found to communicate with young people before they begin to smoke. Past attempts to utilize the interpersonal communications process to encourage non-smoking among adolescents and pre-adolescents has not met with much success. Perhaps, an independence appeal could be better employed as a global value that, for many young people, would have greater salience than the values which smoking behaviors satisfy. An appeal to independence may ultimately establish or reinforce the ideal of being a person who is not enslaved by addiction to nicotine or any other kind of drug. Thus, the campaign may have benefits for other drug control programs as well. The danger of this lumping of other drugs with nicotine would be that the campaign may have the undesired side-effect of lessening the perceived severity of hard drug usage, by lumping hard drugs with the common cigarette.

Since it seems unlikely, at least to these writers, that an anti-diffusion campaign against smoking is likely to affect the seemingly common adolescent sense of immortality, health appeals do not seem to hold much promise for a significant fraction of this segment. With many of the serious health effects of smoking usually taking ten to twenty years to develop, young people are likely to continue to believe either "It can't happen to me," or "I'll quit before then." One appeal that may be meaningful to young people would be that youth is a special time of life, why complicate it with a habit like smoking, a habit that is likely to impede involvement in other activities that are so much a part of being young. Being free to enjoy life, unencumbered by dependency on cigarettes or other drugs may well constitute a much more effective appeal than all the talk, now often regarded as cliches, that smoking is harmful to your health.

Whatever appeals are used, be they independence or others, the goal would be the same; that is, to try to increase the salience of non-smoking as an ideal through interpersonal reinforcement.

#### The Social System

A social milieu change that may weaken the social underpinnings for smoking behaviors would be further restrictions on areas where smoking by children is allowed. If smoking becomes inconvenient or, in other words, a deliberate act, then the nature of the dependence of smokers on their habit will become more obvious with the hoped for eventual benefit of increasing the salience of the independence motive for non-smoking.

It seems that any campaign to restrict smoking areas for adolescents would have to be gradual, at a carefully monitored pace. If the restrictions increase too rapidly, non-compliance would probably be pervasive, weakening, if not destroying, the effectiveness of the campaign. It must be recognized that some hard-core smokers will not quit, regardless of how inconvenient the habit becomes and in spite of whatever social pressures are generated. Time must be allowed for these smokers to adjust gradually their activities in compliance with the increasing restrictions. As their smoking behaviors become gradually isolated from the mainstream of social and educational activity, the perception of smoking as a deviant behavior will gradually increase.

Another danger of too rapid an escalation of smoking area restrictions would be that the appeal of smoking as a forbidden fruit would be increased. The generation of smoking dens, similar to the speakeasys of the Prohibition Era, would not be desirable, though the emergence of such dens may lessen smoking in public places.

#### Time Dimension

Typically, the diffusion of an innovation overtime refers to the rate of penetration of the population at large. It took roughly sixty years before cigarettes entered the growth stage of the product life cycle. It may take at least as many years before a reverse diffusion process will have a significant impact on cigarette sales. This is because the rate of decline as the result of any anti-smoking effort aimed only at prospective smokers will be proportionate to the size of this target group relative to the total smoking population. Campaign objectives, to be realistic, must reflect this effect.

#### Adoption Process

In recognition of the fact that the majority of smokers started before the age of 18 and that the older the individual the less likely smoking will be taken up, one strategy for decreasing cigarette consumption among youths is indirectly to force them to delay the trial of cigarettes (e.g., as might happen with a more vigorous, more extensive athletic program in the schools). And in recognition of the fact that the adoption of cigarettes is rarely preceded by an extended choice process, one strategy is to counteract the influence of the adult and peer smoking models by reducing direct and indirect social pressure. Yet another strategy would be to encourage prospective smokers in this age group to make an overt and deliberate consideration of the pros and cons of smoking before trying even the first cigarette.

#### Summary and Conclusions

On the assumption that people are often too close to the forest to see the trees, the writers have tried to place smoking behaviors in a new perspective. By viewing smoking as innovative behavior, diffusing through successive "generations" of age cohorts, we have attempted to draw on diffusion theory as a conceptual framework for developing insights into alternative approaches for retarding the continuing and accelerating diffusion of cigarette smoking among young people.

The extension of the concept of diffusion, in the context of viewing it as a successive and on-going process, through sequential groups of age cohorts, has potential for applications to other "innovations" as well. Any product, service, or behavior for which the use of or need for is largely restricted by processes of physiological development or by social mores and folkways, to individuals of a certain age, can be meaningfully viewed as the diffusion of an innovation. Examples of other products and services where such age cohort diffusion would occur would include shaving gear, alcoholic beverages, motorized vehicles, cosmetics and contraception.

By recognizing the concept of age cohort diffusion, marketers and others concerned with consumer behavior can draw on the vast body of interdisciplinary research known as diffusion of innovations to gain new insights into the problems that beset products and services subject to age cohort diffusion. Empirical tests of this extension of diffusion theory are planned by these writers, and it is to be hoped that others will test the applicability of age cohort diffusion as well.

#### References

Richard N. Farmer, "Would You Want Your Son to Marry a Marketing Lady?," Journal of Marketing, 41 (January 1977), 15-18.

Phillip Kotler, Marketing Management: Analysis, Planning and Control. (Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1976), 18.

Paul F. Lazarsfeld and Robert K. Merton, "Mass Communication, Popular Taste, and Organized Social Action," in William Schramm ed., Mass Communications, (Urbana, University of Illinois Press, 1949), 459-480.

Joseph C. Robert, The Story of Tobacco in America, (Chapel Hill: University of North Carolina Press, 1969).

Thomas S. Robertson, "The Process of Innovation and The Diffusion of Innovations," Journal of Marketing, 31, (January 1967), 14-19.

Everett M. Rogers, Diffusion of Innovations, (New York: Glencoe Free Press, 1962), 12.

Everett M. Rogers, "New Product Adoption and Diffusion," Journal of Consumer Research, 2, (March 1976), 290-300.

G. D. Wiebe, "Merchandising, Commodities and Citizenship on Television," Public Opinion Quarterly, 15, (Winter 1951-52), 679-691.

Tannis M. Williams, Summary and Implications of Review of Literature Related to Adolescent Smoking, Department of Health, Education, and Welfare, National Clearinghouse for Smoking and Health, Bethesda, MD., 1972.

Gerald Zaltman and Melanie Wallendorf, "Sociology: The Missing Chunk Or How We Missed The Boat," Contemporary Marketing Thought: 1977 Educator's Proceedings of the American Marketing Association, Edited by Barnett A. Greenberg and Danny N. Bellenger, Chicago: American Marketing Association, 1977, 235-38.

Thomas S. Robertson, University of Pennsylvania

## Introduction

These papers present an opportunity to discuss the development of theoretical perspectives within the marketing field. Each paper has taken a somewhat different approach to extending consumer behavior knowledge and the relative merits of each approach will be discussed.

Smoking Behaviors as a Diffusion Process within Age Cohort Groups: An Application of the Societal Marketing Concept

The novelty of this paper by Smith, Olshavsky, and Smith is the borrowing of diffusion theory and applying it to a marketing decision context involving the design of anti-smoking campaigns. The logic of the paper is persuasive and the only concern is a desire for the authors to go further in the following areas:

- Future research would be enriched if the authors could offer specific hypotheses for testing based on the diffusion literature and their knowledge of the smoking literature.
- It would be of further value if they could suggest particular research designs for testing the ideas which they offer.

This paper is based theoretically on an applied problem, as is much theory generation in marketing. What is somewhat different, however, is the direct application of a particular theoretical framework to this applied problem. This is a very "efficient" process for advancing theory development in consumer behavior, but does raise a number of questions, such as:

- Why was this particular theory selected?
- Does this theory adequately represent the complexity of behavior -- is it complete in itself?
- Are multiple theories needed to explain the complexity of the behavior?
- Is this the "best" possible theory for this problem?

Justification and rationale in response to these questions would be helpful for the reader.

Extending Product Life: Technology Isn't the Only Issue

This paper by DeBell and Dardis takes as its central concern the applied problem of extending product life. The authors then address this problem based on survey research concerning consumer purchase and disposal decisions for appliances.

The intriguing part of this paper, in my opinion, is the data on disposal decision-making. The research literature in consumer behavior is so biased toward acquisition and totally ignores the disposal decision process (Jacoby, 1978).

As a pilot study this paper provides data of value to future research and marketing strategy formulation. In particular, however, let me address some comments which may advance further research on this topic.

1. Are there different disposal decision processes by market segment? The present small-scale study is limited to a somewhat older and above average education and income sample. Ferber (1977) has recently commented on the dangers of convenience samples.

2. What theoretical perspectives or concepts might explain the results that the present study obtains?
3. Can this type of research be tied more closely to the burgeoning research which takes an information-processing view to acquisition behavior?
4. Can a situation-specific model of disposal behavior be derived building on the work of Belk (1975), which might help explain differences in behavior by product category.
5. Can future research probe more deeply into the dynamics of family decision-making in disposal decisions?

A Path-Analytic Investigation of Life Satisfaction Among Elderly Consumers

The objective of this study by Bearden, Gustafson and Mason is to test a theoretical explanation of life satisfaction among elderly consumers. As such, the paper succeeds very well. The theory suggested is cohesive and fairly robust and communication with the reader is very efficient since the model is well specified.

The question could be raised as to why this particular model was selected for testing; yet no particular justification is offered. The authors, however, do alert the reader that they are testing one framework and not developing a complete explanation of elderly life satisfaction. There are also some measurement questions which could be addressed in terms of the validity and reliability of the scales, despite their apparent logic.

## Conclusion

These papers illustrate different approaches to theoretical development including: (1) borrowing theory and applying it to a new problem; (2) starting with the problem and conceptualizing research which will appropriately bring empirical data to bear on that problem; and (3) testing a particular conceptual model appropriate to a research issue.

Given the current state of affairs in consumer behavior theory development, we need to pursue these approaches simultaneously. However, much of the research to date fails to build toward cohesive theory. As Robertson and Ward observed in 1973, "Since many different concepts from diverse behavioral science areas have been applied separately to marketing problems, findings are disjointed and lack consistency and integration in formal, predictive theories and models" (p. 14). Jacoby (1978) has also questioned the value of much consumer behavior research due to the lack of theory and the lack of appropriate methodology and analysis.

Thus, what the consumer behavior field badly needs is integration of concepts and the development of cohesive models appropriate to particular problems. In this sense, the paper by Bearden, Gustafson and Mason is particularly refreshing. It is unlikely that we can continue to borrow without restraints from the behavioral sciences in general and it is grossly inefficient to conduct idiosyncratic research free of theoretical constructs. Perhaps some progress is being made in

building more cohesive conceptual foundations.

#### References

Russell W. Belk, "Situational Variables and Consumer Behavior," Journal of Consumer Research, 2 (December 1975), 157-164.

Robert Ferber, "Research by Convenience," Journal of Consumer Research, 4 (June 1977), 57-58.

Jacob Jacoby, "Consumer Research: A State of the Art Review," Journal of Marketing, 42 (April 1978), 87-96.

Thomas S. Robertson and Scott Ward, "Consumer Behavior Research: Promise and Prospects," in Consumer Behavior: Theoretical Sources, eds. Scott Ward and Thomas S. Robertson. Englewood Cliffs, NJ.: Prentice-Hall, Inc., 1973.

DISCUSSION PAPER: SOCIOLOGICAL PERSPECTIVES

Gerald Zaltman, University of Pittsburgh

Smoking Behaviors as a Diffusion Process Within  
Age Cohort Groups: An Application of the  
Societal Marketing Concept

This is an intriguing paper. Smith, Olshavsky and Smith take several concepts commonly used to understand the process whereby innovations spread, and ask whether these concepts might also be used to prevent that process from occurring. Their answer is a tentative "yes." Their approach is a definite must, but with a major qualification. Before addressing this qualification and a few other concerns, let me identify some especially strong features of this paper.

Although the authors nowhere formally state a sociological perspective on life-style, it is implicit in many of their ideas. Among other things, this perspective describes life-style as a group phenomenon, representing shared values, linking diverse activities in people's lives, and involving a continued process of adoption, emulation, and change. Life-style is an important concept in the health field, but health education largely ignored the important dimension of a sociological perspective. Smith, Olshavsky and Smith demonstrate very nicely the practical implications of adopting this perspective.

The idea of gradually engineering norms so that smoking is perceived as deviant social behavior is a potentially powerful notion. Nearly all health education efforts focus on the harmful physiological effects of smoking. It does not address or correct the initial causes of smoking. The authors raise a significant point in suggesting that the social causes of smoking be addressed by anti-smoking campaigns as well as the harmful physiological consequences.

The authors also distinguish between encouraging people to cease smoking and not starting at all. This important distinction could have been treated more systematically, but it is still present and is an important contribution. Through the tactics they suggest, the authors recognize that it is one task to get people to stop something they've already started, and quite another task to get people to not start in the first place. Many anti-smoking campaigns I've examined have failed to make this important distinction.

My major reservations about the application of diffusion theory to anti-diffusion efforts is that, while the basic concepts may be relevant, the way in which they traditionally function may not be relevant. Diffusing resistance to change is not necessarily the same as diffusing acceptance of change, just as learning and forgetting are not the same. For example, the authors discuss how the conventionally noted attributes of innovations may be considered in designing an anti-smoking campaign. Actually, these ideas have been used rather widely in this context and in other health-related contexts. They have also been largely ineffective. This does not mean that the concept of innovation attributes as the authors use it is not relevant. Perhaps attributes relevant to the acceptance of an idea or practice are not those which are most relevant for the rejection of an idea or practice.

There are, of course, a number of specific points in this paper that one could take issue with. These would range from the quite erroneous suggestion that no real decision process is involved in the decision to take up smoking to the erroneous authorship attribution of the major elements of the diffusion process (Katz, et al., not Rogers) to an apparent lack of awareness of recent work related to anti-smoking. However, overall the paper is provocative and very worthwhile reading for its practical ideas for reducing the prevalence of smoking.

Extending Product Life: Technology  
Isn't the Only Issue

The major contribution of this paper is the focusing of attention on product disposal decisions. This important type of decisions has received very little attention and DeBell and Dardis are to be lauded for venturing forth in this area. Their data confirm their commonsense assumption that differences in disposal behavior exist between products having mechanical/performance obsolescence and those having fashion/technological obsolescence. As they point out in their discussion, the confirmation of this assumption has implications for the allocation of efforts to improve product durability. They also raise the interesting issue as to whether durability improvement efforts should be expended at all in areas where disposal decisions are largely influenced by fashion/technological obsolescence.

The paper is admittedly a pilot study. As such it adequately establishes the case that disposal decisions may not be the same as acquisition decisions, and that disposal decisions may vary according to type of product. In themselves these are not exciting ideas but are necessary to establish empirically prior to pursuing their more interesting implications. I would urge the authors and others to direct subsequent efforts to a deeper understanding of such issues as the following:

1. Why fashion/technological obsolescence is more salient for some products and not for others? Perhaps more importantly, are there differences for a given product in the kind of individuals and families who do and do not replace that product for fashion/technological reasons?
2. Are there sociological and psychological differences among individuals and/or families in their disposal decisions? Family decision-making processes with regard to disposal decisions would be an exciting avenue of research.
3. Much more research comparing acquisition and disposal decisions is needed. One could speculate endlessly about exciting relevant points of comparison such as who the relative influentials (in, say, the family) are in each process?
4. Additionally, how (if at all) are acquisition and disposal decisions interrelated? If disposal also involves replacement, the two types of decisions are unlikely to be independent of one another. This becomes especially intriguing if each spouse is differentially influential in the two types of decisions.

5. Carrying the issue of interrelatedness between acquisition and disposal decisions further, we might want to question or examine more closely the assumption that these are two separate decision processes. For some product or service categories the interdependence may be very great. Incidentally, the Smith, et al. paper discussing smoking and anti-smoking is clearly an example where the "acquisition" decision and "disposal" decision may be quite independent.

A Path-Analytic Investigation of Life  
Satisfaction Among Elderly Consumers

The purpose of this paper was to predict and explain overall elderly satisfaction with life. This is an important issue and a very interesting paper for persons concerned with this issue. However, the audience for this paper is primarily concerned with consumer behavior and, indeed, one of the explanatory variables involves consumer satisfaction. My comments will be divided into two parts. The first part concerns consumer satisfaction as an explanatory variable. The second part concerns issues raised by the methodology of the paper.

It is refreshing to see the authors treat consumer satisfaction as an independent variable and overall life satisfaction as a dependent variable. From a policy standpoint, one of the major reasons for studying consumer satisfaction is precisely because it should be an important contributing factor to life satisfaction in general. This paper challenges this widespread notion. The direct effect is neither statistically nor substantively significant. (The zero order correlation is not substantively significant.) Moreover, the indirect effect of other variables through consumer satisfaction is one of the weaker total effects. Unfortunately, the authors do not pay attention to this very important challenge to existing assumptions. Their failure to address this important matter is the principal weakness of the paper.

The model the authors test is robust. They appear a bit apologetic for explaining only thirty percent of the variance in the dependent variable. This apologetic posture would be warranted if the nature of the phenomena under study were such that there lurked a seventh or eighth independent variable which, if added to the model, would account for a large additional percentage of the unexplained variance. My own sense is that the dependent variable is sufficiently complex and multifaceted -- as most interesting independent variables are -- that a large number of additional explanatory variables would be needed to explain substantially more variance. Moreover, none of these additional variables alone would have an explanatory contribution greater than those used by the authors. Thus I feel the authors have developed and tested a good model which they admit is only a part of the total reality. The approach they follow is one I recommend strongly to others.

AN EXPLORATORY STUDY OF  
TODDLERS' SATISFACTION WITH THEIR  
TOY ENVIRONMENTS

Gail Bjorklund, Hood College  
Richard Bjorklund, George Mason University

Abstract

Investigation of cognitive aspects of consumer socialization among children can be helpful in identifying early experiences which influence later consumer behavior patterns. An exploratory study was conducted to determine how variations in the toy environment according to quantity and qualitative category would effect satisfaction measures among toddlers. The findings showed promise for future research of this nature.

Introduction

Consumer socialization has been viewed as a process whereby persons learn to evaluate information in a choice situation (Ward). As a result, the consumer socialization process depends on information-processing skills. Since information processing skills vary according to ones level of cognitive development (Ward), cognitive development is also an important aspect of the consumer socialization process. Consequently, research on the cognitive aspects of consumer socialization can further our understanding of ways to improve consumer behavior by furthering information-processing skills.

The consumer socialization process can be assumed to begin during infancy (from birth to 1 year) and progress through toddlerhood (from age 1 to 2½) and continue throughout life. At the toddler stage, the toys parents purchase for toddler consumption may have a significant effect on cognitive development. For toddlers, the object environment, of which toys are an important component, represents an important source of stimulation for cognitive growth (Piaget and Inhelder, 1964). If a child's behavior indicates he prefers one toy environment over another, there is reason to believe that the more satisfying toy environment will have a positive effect on the consumer socialization process.

Nevertheless, little information is available to parents on what toy environments are most suited for different age groups of children. Therefore one must assume parents are purchasing toys based on their perceptions of which toys satisfy their children. The purchase decisions are not being made based on information obtained from objective studies of toddler needs. Similarly, the lack of published research leads one to question the criteria toy companies employ to position their products for specific age groups. If toy companies have conducted unpublished research aimed at developing and positioning toys that satisfy child needs, these efforts should be applauded. However, if toy companies are marketing toys for different age groups based on criteria other than those developed from objective, child related research, clearly they too would benefit from such research.

The objectives of the present research are to identify selected aspects of the toy environment that appear to be most satisfying to toddlers and to determine if age differences correspond to these differing toy environments. First, the traditional measures of satisfaction are examined from the marketing literature. Then the literature on sensorimotor cognitive development and environmental stimulation are used to define the toy environment variables and to provide behavioral measures

of satisfaction. Finally, an experiment is conducted that attempts to (1) identify toy environments that satisfy toddlers and (2) determine if a "matching" of toy environments with levels of cognitive functioning exists.

Past Research

Satisfaction

Satisfaction might be defined as the fulfillment of a need or want (Aiello, in Day, 1977). If a product fulfills a need or want presumably more of the product will be consumed. Recently Day (1977) and Pfaff (in Hunt, 1977) have discussed alternative approaches to conceptualizing and measuring satisfaction.

The first approach, or utility approach, is related to the economists concept of satisfaction. Measurement of satisfaction is generally accomplished by analyzing respondent generated preference distributions (Day, 1965, Srinivasan, 1975) ideal points (Green, 1975a) or conjoint measurement determined scales (Green, 1975b). From this data indices of consumer satisfaction can be developed.

The second approach to satisfaction has focused on the extent to which prior expectations regarding product performance have been confirmed or disconfirmed in the consumption process. Here, measurement of the consumer's satisfaction with product related attributes takes place after a consumption experience. The consumer is generally asked to respond to a simple satisfaction/dissatisfaction scale. The Howard and Sheth (1969) model of consumer behavior approaches satisfaction similarly.

"Satisfaction is defined as the buyers cognitive state of being adequately or inadequately rewarded in a buying situation for the sacrifice he has undergone. The adequacy is a consequence of matching actual past and purchase experience with the reward that was expected from the brand." (1969)

Neither of the traditional approaches to satisfaction are directly applicable in this study. The immaturity of toddler subjects prohibits the use of prior expectancy since subjects would have little or no experience upon which to base their expectancies. In addition the expectancy and utility approach requires verbal or written responses which cannot be obtained from infants. However, measures of satisfaction can be based on observable behavioral data. It seems reasonable to assume that infants would most directly express their degree of satisfaction/dissatisfaction through some overt behavior. Furthermore, since their consumption experience is limited and their learning process just beginning, a behavioral measurement might be somewhat less confounded by intervening psychological variables than if adults were the subjects.

The development of appropriate measures of a toddler's satisfaction with his toy environment can find little guidance in the presently available marketing literature. The scales used have relied on verbal and written measures of avowed satisfaction and the products evaluated have not included characteristics of the toy environment.



Thus, new measures of consumer satisfaction must be designed which attend to those characteristics of the toddler's play with toys which most directly relate to satisfaction. In addition, the attributes of the toy environment which are likely to be most sensitive to changes in satisfaction must be ascertained.

Two areas within the developmental literature can provide relevant information towards the design of the present study - the literature on sensorimotor cognitive development and the literature on environmental stimulation.

#### Sensorimotor Cognitive Development

A cognitive theory of consumer motivation and behavior attends to phenomena such as knowing, understanding, perception and information processing (Markin, 1977). These cognitive processes, according to Markin, serve two basic functions:

- 1) they are purposive in that they serve the individual in his attempt to achieve satisfaction of needs.
- 2) they are regulatory in that they in large measure determine the direction taken in the consumer's attempt to attain satisfaction of his initiating needs." (p. 38, Markin, 1977)

Thus, there exists a dynamic interaction between the individual and the environment as "sensory input is transformed, reduced, elaborated, stored, recovered and used" (p. 43, Neisser, 1967).

Cognitive activity during the first two years of life (infancy) is based upon direct sensory and motor forms of activity and thereby is defined as the "sensorimotor period" by Piaget (1952). Several forces influence cognitive development, including the infant's interactions with the physical environment. During the sensorimotor period the primary form of interaction with the physical environment occurs during play with objects such as toys. Sensorimotor interactions with objects become an important vehicle through which the child derives satisfaction of both physiological and psychological needs, and thus the continuing motivation to interact with the environment. The interrelationships between needs, satisfaction and cognitive activity have been described by Piaget as follows:

"The need sets in motion the act and its functioning, but this functioning itself engenders a greater need which from the very first goes beyond the pure satisfaction of the initial need. Hence, the primary fact is neither the need anterior to the act nor repetition, but it is the total relation of the need to the satisfaction" (p. 170, 1952).

Piaget does not indicate how specific environmental conditions can influence variations in satisfaction or cognitive functioning. However, he does imply that children with a richer variety of concrete experiences with the object world will be equipped with a more complex repertoire of exploratory behaviors to apply to their object interactions (Piaget, 1952). It seems reasonable to assume that a richer, more complex level of cognitive functioning will result in a better ability to differentiate satisfying experiences. In order to determine how satisfaction is influenced, a set of operational measures of "richer," "more complex" levels of cognitive functioning must be determined.

Kagan's investigation (1971) of attention among infants as it related to differences in what was termed the "play tempo" provides a useful precedent for measuring levels of cognitive functioning during infant's play with toys. Differences in play tempo could be distinguished as "reflective" or "impulsive" based upon the measures of the time spent with a toy and the number

of acts performed. Impulsive children used fewer acts per time unit and spent less time in each toy contact, while reflective children had opposite play styles. These differences were stable over time and directly related to later cognitive performance.

Thus, it appears that longer periods of sustained contacts with toys and greater numbers of acts are indicative of an increased sense of effectiveness. More needs are satisfied as more exploration occurs. Therefore, the satisfaction that a child derives from certain levels of toy quantity and toy qualitative categories should be measurable by behavioral variables of acts and contact time.

#### Environmental Stimulation

Awareness of the importance of infancy toward long-term cognitive development has resulted in much attention being placed on optimizing early environmental stimulation, especially during the more "critical period" between 10-18 months (White, 1973). The importance of infant-object interactions during the later half of sensorimotor development has resulted in particular attention being focused on toys as stimulus objects. The ability of environments to vary in their ability to stimulate cognitive activity, and thereby increase satisfaction, can be related to investigations of differing toy environments and their impact on cognitive performance.

White and Held (1966) were early investigators of the potential for accelerating the onset of sensorimotor skills through an enriched toy environment. Their results implied that toys would stimulate developmental gains upon the onset of prehension skills. Subsequent longitudinal studies by Kagan (1971), White (1973), Wachs (1976) and Yarrow (1975) have indicated that stimulation within the home environments during the first two years could have significant impact on later cognitive achievement. Toys were found to be significant stimulation variables within these home environments.

The results of these and other investigations of early environmental stimulation have consistently isolated several significant characteristics of toy environments. Wach's (1976) analysis revealed the significance of

- 1) the predictability or regularity of the home environment, including toys
- 2) the adequacy of stimulation as defined by such criteria as the amount of auditory visual and tactile stimulation and the variety of stimulation
- 3) the intensity of stimulation as determined by such factors as the number of toys, and
- 4) availability of toys offering feedback which related to the child's present level of cognitive functioning.

Similarly, Yarrow, Rubenstein and Pedersen (1975) found that varied toy stimuli and toy responsiveness had strong relationships with developmental outcomes. McCall (1974) also discovered that children self-selected those toys which "matched" their level of development.

Thus, it appears that toy quantity has the potential for influencing cognitive performance. Prior research, however, has not compared the effects of specific toy quantity levels under controlled experimental conditions.

In addition, there seems to be an important relationship between a child's age and the characteristics of the toys themselves. Hunt (1961) was the first to postulate the importance of "matching" the environment to a child's cognitive abilities and the research of Wachs (1976) and McCall (1974) tend to substantiate the hypothesis.

However, specific categories of toys have not been clearly defined in the past research. An obvious source of information for determining appropriate categories of toys is the work of Piaget and others on the progression of infant-object interactions during infancy.

Piaget (1952) has indicated that the sensorimotor period can be decomposed into six stages of cognitive development. In infants progress through these six stages, qualitative differences can be observed in exploration of objects during play. Children within the later three stages (from about eight to twenty-four months), objects such as toys assume a greater role in the child's cognitive activity, and thereby provide a major source of satisfaction.

The actual progression of infant object interactions with toys during the later three stages of the sensorimotor period have been empirically confirmed by Sinclair (1970), Inhelder (1971), Fenson (1976), and McCall (1974). Three major classes of behavior have been consistently identified:

- 1) Simple, sensorimotor exploratory acts such as shaking, mouthing, moving.
- 2) Relational acts where objects have been simultaneously manipulated in ways that introduce relationships and/or arrangement.
- 3) Symbolic acts where objects have been used in a make-believe or representational manner.

The present study has selected toy category definitions which correspond to the above classes of behavior to further test the potential for enhancing satisfaction by providing toys which correspond to a child's cognitive stage.

#### The Present Study

The purpose of the present study is to extend the limited research on the effects of the quantity and quality of toys on the satisfaction of toddlers. Child development researchers have theorized the importance of play interactions with toys and satisfaction research has received recent attention in the marketing literature. Nevertheless, little, if any, empirical research on the satisfaction of toddlers with their toys has taken place in either the marketing or child development literature. This study is unique in that it investigates toddlers' satisfaction with their own toy environment using behavioral measures of toy satisfaction.

More specifically, this study tested the following hypotheses:

- H1: Differences in toddler's satisfaction with different toy quantity conditions will exist.
- H2: Differences in toddler's satisfaction with different toy categories will exist, and these differences will correspond to the "matching" age group.

#### Methodology

A split plot factorial design was employed in this study where the quality and quantity of toy were experimental conditions. The experimental design was blocked on the age and sex variables.

#### Experimental Conditions

The quality of toys was defined based on the three categories of play behavior reported by Sinclair (1970), Inhelder (1971), McCall (1974) and Fenson (1976).

The definitions of these three categories are as follows:

- (1) Responsive toys: Objects designed to emit auditory and/or visual responses when manipulated by a child. For example, a squeak toy would produce a sound when touched or squeezed.
- (2) Organizational toys: Objects designed for the orderly arrangement of related parts such that the child's manipulation can introduce relationships among the objects' parts. For example, a set of stack rings can be taken apart and replaced on the stick in order of size.
- (3) Symbolic toys: Objects designed to be used in make believe or "as if" the objects were something else. For example, making believe there is someone to talk to on a toy telephone, or using a doll as if it were a real baby.

Toy quantity levels of 3 toys, 12 toys and 21 toys were chosen to represent low, medium and high intensities of stimulation. Pretesting revealed noticeable differences in toddler's play behavior with these quantity levels. At each quantity condition, an equal number of toys from each of the three toy categories were randomly selected.

Measures of toddler satisfaction with toys were defined as the time per toy contact and the number of acts per toy contact. Satisfaction was assumed to vary directly with both measures. These measures defined as follows:

Time Per Toy Contact: Once a child contacted a toy, the duration of the contact was measured in seconds across all contacts with the toy and then averaged for the total number of times the toy was contacted during the ten minute testing period. This measure was calculated over all toy category and toy quantity levels.

Acts Per Toy Contact: An act was defined as (1) a single form of behavior such that the child's actions could be described by a single verb or verb phrase expressing manipulation and (2) having an observable beginning and ending. Cues used to operationalize this definition included a change in the form of behavior, a pause of more than five seconds separating acts, or the use of a different toy in performing the same act. The number of acts observed was summed for each toy contacted and averaged across the total number of contacts with that toy.

#### Subjects

Due to the distinct but rapid developmental period of infants (Piaget, 1952) and the importance of family background influences (Kagan, 1971; Yarrow, 1975; Wachs, 1976) subjects were carefully selected. Children in the age groups of 12 months, 16 months, 20 months and whose parents were college educated and professionally employed were selected for this study. A convenient sample of 24 infants was obtained from the Bloomington, Indiana area. The sample size consisted of 4 males and 4 females for each age group.

Random assignment of subjects to toy quantity levels would have produced cell sizes too small for analysis. To avoid this problem repeated measures on subjects across toy quantity levels were employed.

#### Experimental Design

The nature of the treatment conditions and the use of repeated measures data suggested the use of a split plot factorial design.

Due to the repeated measures nature of the design there are two potential limitations of this experiment:

- (1) The effect of treatment order can influence the measurement of the dependent variables due to carry over effects from prior treatment.
- (2) There is a potential for correlation among scores within treatment levels because the same subjects are observed under all treatment conditions.

The first limitation is minimized by randomizing the order of quantity level presented to each subject. The second limitation does not appear relevant. Subjects in this study have undeveloped capacity for long term memory and thus a small likelihood of remembering earlier treatment sessions. Therefore the potential for correlation among scores in the within subject effects should be small; thus the assumptions for univariate repeated measures designs are closely approximated.

#### Toy Selection

Thirty-five toys were initially selected for this study. Most toys came from major toy manufacturers and were recommended for children from the age group under study. The remaining toys were objects commonly found in households which children might use as toys (i.e., a comb, a plastic cup and saucer).

Three experts, knowledgeable of the study's objectives and trained in the definitions of the three toy categories, independently classified the thirty-five toys into the three toy categories. The experts' classifications were reviewed and ambiguous toys, where interrater classifications did not agree, were eliminated. Toys for the 21 toy level were randomly chosen from the final set of 30 toys subject to having seven toys in each toy category. Toys for the 3 and 12 toy levels were randomly chosen from the 21 toy level.

#### Data Collection

The laboratory for this experiment consisted of a carpeted room containing a table and chair for a parent and a hidden video tape equipment. Parent presence was maintained to help the child adapt to the experimental environment.

During the adjustment to experimental environment period, each parent was told to interact with the infant to facilitate quick adjustment. No infant adjustment problems occurred. When it appeared the child was comfortable with the environment the 10 minute testing period commenced. Parents were instructed not to influence their child's selection of toys; thus, the parents were told to use only general statements like "Look at the toys" or "Play with the toys" when talking to their child. Video tape equipment was used to record all testing periods.

During each testing period a child was randomly assigned to a quantity level and the toys within the toy categories were randomly selected. Each child participated in three testing periods, corresponding to the three quantity levels. The average time between testing periods was three days.

Two raters who had been thoroughly trained in the study's objectives and the definition of acts and contacts coded each video taped session. Prior to actual testing, but after the rater training period, pretest subjects were used to check interrater reliability. The number of acts per toy category recorded by the first rater was correlated with the number of acts per toy category recorded by the second rater. An average correlation coefficient of .92 ( $p < .05$ ) was obtained.

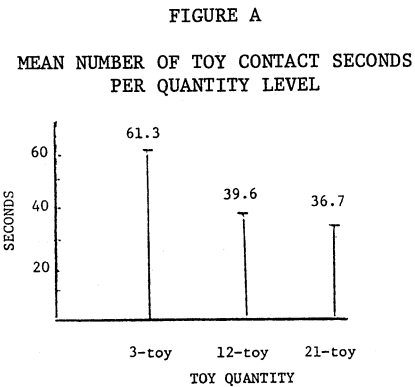
## Results

The hypotheses related to the effects of toy quantity and toy category on satisfaction were tested by two measures: time per contact and acts per contact. The results of analysis of variance on the two measures are presented, followed by a discussion of the tested hypotheses.

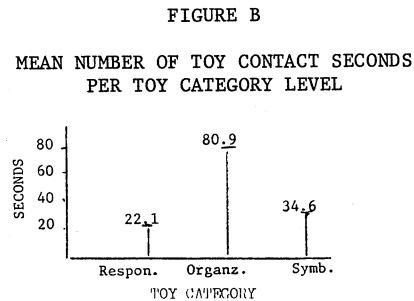
#### Time Per Toy Contact

The analysis of time per toy contact seeks to identify the toy quantity and quality treatment conditions that relate significantly to longer periods of individual toy contact, as compared to those conditions in which a child tends to move quickly from one toy to another without sustaining prolonged toy contact.

**Main Effects.** There were significant differences in the time of each toy contact at different toy quantity levels ( $F = 4.7572$ ,  $df = 2,36$ ,  $p < .05$ ). When only three toys were present, each toy contact was significantly longer than at either the 12 or 21-toy levels. A Tukey test of multiple comparisons indicated that there was no significant differences in the duration of each toy contact between the 12 and 21-toy levels. Figure A shows the average time per contact for each child at the three toy quantity levels.



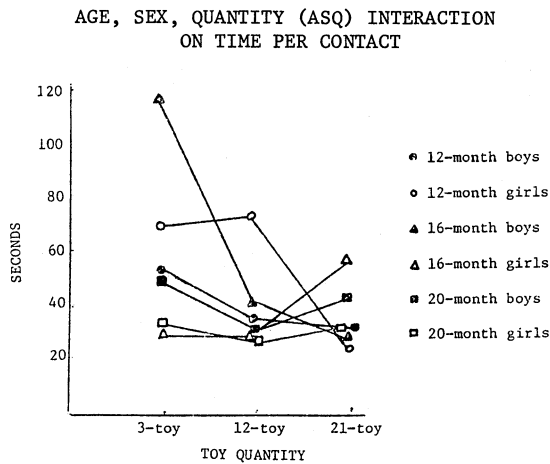
The main effect of toy category (C) had particular influence on the duration of each toy contact ( $F = 20.0355$ ,  $df = 2,36$ ,  $p < .01$ ). A Tukey test of multiple comparisons showed that all children remained in contact with a toy of the organizational category significantly longer than with either the responsive or symbolic category, as indicated in Figure B.



**Interactions.** The age, sex, and toy quantity (ASQ) interaction ( $F = 3.3619$ ,  $df = 4,36$ ,  $p < .05$ ) seems to reflect age and sex differences at the 3-toy level. Figure C shows that 16-month old boys varied the most dramatically, remaining 123 seconds with a toy at the 3-toy level. This dropped to 43 seconds and 27 seconds respectively at the 12 and 21-toy levels. Twelve-month

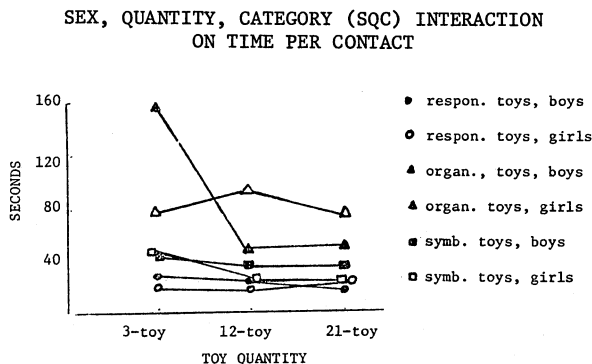
old girls contributed modestly to the interaction, spending more time per toy in the 3 and 12-toy conditions when compared with the time spent by older girls, by boys at 12 and 20-months of age, or with more toys present.

FIGURE C



A significant sex, quantity, and toy category (SQC) interaction ( $F = 3.5782$ ,  $df = 4,72$ ,  $p < .05$ ) is illustrated in Figure D. It appears that when there were only three toys to choose from, boys spent the longest amount of time with the organizational category of toy.

FIGURE D



Acts Per Toy Contact

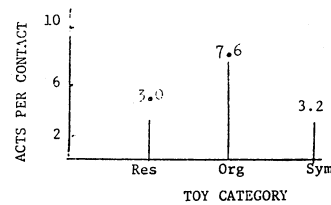
The mean number of acts performed during a single toy contact provides additional information about the effects of toy quantity and category differences on satisfaction. For example, an infant may have performed more total acts at higher toy quantity levels. However, if infant was not "satisfied" by the toy condition, these acts may have been performed while quickly moving from one to another resulting in fewer acts within each toy contact. On the other hand, a child who was satisfied with the toy environment would remain in continuing contact with a toy and perform many acts within this single contact period. In order to identify how toy quantity and category conditions relate to satisfaction, an analysis of the number of acts per toy contact was carried out.

Main Effects. Significant differences were identified for toy category (C). During each contact with an

organizational toy, children performed significantly more acts ( $F = 24.3929$ ,  $df = 2, 36$ ,  $p < .01$ ). Figure E illustrates the mean number of acts each child performed within a single contact for each toy category.

FIGURE E

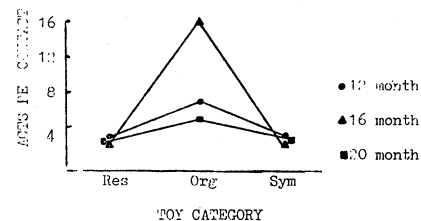
MEAN NUMBER OF ACTS PER TOY CATEGORY LEVEL



Interactions. Children's acts per contact with various toy categories did vary according to age, as evidenced by the age and toy category interaction ( $F = 3.8984$ ,  $df = 4,36$ ,  $p < .05$ ). While children of each age group used more acts per contact with organizational toys, children at 16 months showed a significantly greater rise, as shown in Figure F.

FIGURE F

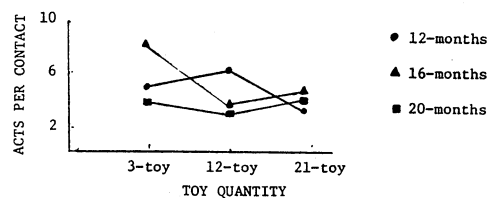
AGE, TOY CATEGORY INTERACTION FOR ACTS PER CONTACT



Acts per contact varied according to age and toy quantity ( $F = 2.8276$ ,  $df = 4,36$ ,  $p < .05$ ). Figure G shows that the performance of 12 month and 16 month old children varied most dramatically: 12 month old children used more acts per contact as toys increased from 3 to 12 and then declined at the 21 toy level, whereas 16 month children decreased their acts per contact as toys increased from 3 to 12 and increased their acts per contact as toys increased to 21. The performance of 20 month old children was not strongly affected by toy quantity.

FIGURE G

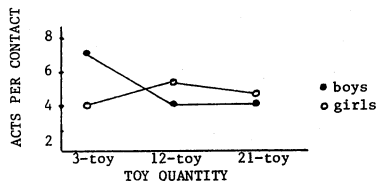
AGE, QUANTITY (AQ) INTERACTION FOR ACTS PER CONTACT



Sex also interacted with toy quantity ( $F = 4.6234$ ,  $df = 2,36$ ,  $p < .05$ ) such that acts per contact among boys declined sharply at the 12 and 21 toy levels as shown in Figure H. Girls, on the other hand, showed a slight increase in acts per contact at the 12 and 21 toy levels.

FIGURE H

SEX, QUANTITY (SQ) INTERACTION FOR ACTS PER CONTACT



The significance of additional interactions of the analysis of acts per contact further demonstrate the various influence of age and/or sex on acts per contact at different toy quality and quantity conditions. Constraints on the length of this paper precludes examination of each interaction; however, several general patterns can be noted. Boys at 16 months tended to use the greatest number of acts per contact. In addition, their acts per contact were significantly greater than all girls and boys at 12 and 20 months when only 3 toys were present and when they were contacting organizational toys.

#### Hypothesis Testing

The above analyses of the measures of contact time and acts per contact were used to distinguish differences in satisfaction based on the effects of toy quantity ( $H_1$ ) and toy quality ( $H_2$ ).

$H_1$  stated that toddler's satisfaction with toys would vary at different toy quantity conditions. The two measures of time per contact and acts per contact differed for the main effect of toy quantity: only the measure of time per contact was effected by toy quantity such that children had significantly longer contact times at the 3 toy level. However, when age and sex differences are considered for each measure, a consistent pattern of higher levels of satisfaction among 16 month old boys at the 3 toy level was evident.

$H_2$  stated that satisfaction would vary with different toy categories, and that these differences would correspond to the "matching" age group. Significant differences in both measures of satisfaction revealed that toddlers in each age group derived greater satisfaction from organizational toys. No age relationships with toy categories were found for time per contact and and for acts per contact; the only relationship revealed was between 16-month old children and organizational toys. Thus, while satisfaction does vary with toy categories, definitive age relationships for each toy category do not seem to exist.

#### Conclusions

An exploratory investigation, such as this, of the cognitive development of toddlers can further our understanding and establish an interest for future research on early information - processing and the consumer socialization process. Such research may also be helpful in identifying those early experiences which influence later consumer behavior patterns as well as later cognitive functioning.

The ability to differentiate more satisfying toy environments appears to be more appropriately measured by behavioral data among toddlers who cannot respond to traditional verbal or written methods. Thus, the measures of time per contact and acts per contact which had been used in prior research on cognitive development, were employed in the present study as measures of satisfaction.

These measures were used to determine how satisfaction among toddlers in three selected age groups would differ according to two common variations in the toy environment - toy quantity and toy category.

The findings indicate that certain differences in satisfaction exist under different toy quantity levels. However, the results for the variables of time per contact and acts per contact were not consistent, making it difficult to form definite conclusions. For example, the 3-toy level was satisfying to toddlers based upon time per contact. No toy quantity main effect, however, was found for acts per contact. Instead, age and sex interacted with toy quantity such that 16 month old boys seemed to derive the greatest satisfaction from the 3-toy level.

One explanation for the variation in the two measures of satisfaction could be related to individual differences in children's prior toy quantity exposure in their home environments. Children may become accustomed to the toy quantity levels generally available, with the result that satisfaction is differentially affected by varying quantity conditions in accordance with prior experiences. Historical data on children's home toy environment, longitudinal studies in controlled quantity conditions, or cross-cultural methods might be appropriate future alternatives for exploring toy quantity effects. In addition, the variation in the toy quantity levels may require the utilization of a more controlled familiarization period to better equate the toy explorations children employ under each quantity condition.

The additional finding that no significant differences existed between the 12 and 21 toy levels for either time or acts per toy contact has important implications. Although children did not appear to react negatively to increasing quantity to 21 toys, neither did they appear to change their play behavior. Therefore, the value of providing a great number of toys is highly questionable, based on the present results.

Differences in satisfaction with different toy categories were found. The organization toy category was consistently associated with longer contact times and greater numbers of acts. Finally, using acts per contact as the measure of satisfaction revealed an a priori expected matching of 16 month old children with organizational toys. In the future research, perhaps a wider range of age groups should be used to test for the matching of other toy categories with age groups. Although all three age groups in the present study preferred organizational toys, younger ages, for example, may show equal satisfaction with all toys while older children may show increasing satisfaction with symbolic toys.

#### References

- A. Aiello, J. Czepiel, and L. Rosenberg, "Scaling the Heights of Consumer Satisfaction," in R. L. Day (ed.), Consumer Satisfaction, Dissatisfaction and Complaining Behavior, Bloomington, Indiana: Foundation for the School of Business, 1977.

- R. E. Anderson, "Consumer Dissatisfaction: The effect of disconfirmed expectancy on Performance." Journal of Marketing Research. (1973), 10, 38-44.
- R. L. Day, Consumer Satisfaction, Dissatisfaction, and Complaining Behavior. Bloomington, Indiana: Foundation for the School of Business, 1977.
- Ralph L. Day, "Systematic Paired Comparisons in Preference Analysis" Journal of Marketing Research, 2 (November 1965) pp. 406-412.
- R. Day and L. Landon, "Toward a Theory of Consumer Complaining Behavior," in A. G. Woodside, J. Sheth, P. Bennett, (eds.), Consumer and Industrial Buyer Behavior New York: North-Holland, 1977.
- L. Fenson & J. Kagan, et al., "The development progression of manipulative play in the first two years," Child Development, 1976, 47, 232-236.
- Paul E. Green, "Marketing Applications of MDS: Assessment and Outlook" Journal of Marketing, 39 (January 1975), pp. 624-31.
- Paul E. Green and Yoram Wince, "New Ways to Measure Consumers' Judgments" Harvard Business Review, 53 (July-August 1975) pp. 107-117.
- J. A. Howard, Consumer Behavior: Application of Theory New York: McGraw Hill, 1977.
- J. A. Howard and J. N. Sheth, The Theory of Buyer Behavior New York: Wiley, 1969.
- J. McV. Hunt, Intelligence and Experience, New York: Ronald Press, 1961.
- B. Inhelder, "The Sensory-Motor Origins of Knowledge," in Walcher & D. L. Peters (eds.) Early Childhood: The Development of Self-Regulatory Mechanisms, New York: Academic Press, 1971.
- J. Kagan, Change and Continuity in Infancy, New York: John Wiley and Sons, 1971.
- R. Markin, "Motivation in Buyer Behavior Theory: From Mechanism to Cognition." in A. G. Woodside, J. Sheth and P. Bennett (eds.) Consumer and Industrial Buyer Behavior, New York: North-Holland, 1977.
- Robert B. McCall, "Exploratory manipulation and play in the human infant," Monographs of the Society for Research in Child Development, 1974, 39, No. 2.
- U. Neisser, Cognitive Psychology, New York: Appleton-Century-Crofts, 1967.
- M. Pfaff, "The Index of Consumer Satisfaction," in H. Keith Hund (ed.), The Conceptualization of Consumer Satisfaction and Dissatisfaction, Cambridge, Mass: Marketing Science Institute, 1977.
- Jean Piaget, Origins of Intelligence in Children, New York: W. W. Norton and Company, 1952.
- Jean Piaget and B. Inhelder, The Psychology of the Child. New York: Basic Books, 1964.
- H. Sinclair, "The transition from sensory-motor behavior to symbolic activity," Interchange, 1970, 1, No. 3.
- V. Srinivasan, "A General Procedure for Estimating Consumer Preference Distributions," Journal of Marketing Research 11 (November 1975) pp. (377-389).
- T. D. Wachs, "Utilization of a Piagetian approach in the investigation of early experience effects: a research strategy and some illustrative data," Merrill-Palmer Quarterly, 1976, 22, 11-30.
- S. Ward, D. Wackman, T. Wartella, "Children and Commercials: Cognitive Development Theory" (paper based on a draft of Chapter 1 from Children Learning to Buy: The Development of Consumer Information Processing Skills, Sage Publications).
- B. White and R. Held, "Plasticity of sensorimotor development in the human infant," in Rosenblith, J.F. and Allinsmith, W., (eds.), Causes of Behavior: Readings in Child Development and Educational Psychology (2nd ed.) Boston, Mass: Allyn and Bacon, 1966.
- B. White and J. C. Watts, Experience and Environment, Englewood Cliffs, New Jersey: Prentice Hall, 1973.
- L. J. Yarrow, J. L. Rubenstein, and F. A. Pederson, Infant and Environment, Washington, D.C.: Hemisphere Publishing Co., 1975.

CHILDREN'S ATTITUDES TOWARD ADVERTISING ON TELEVISION AND RADIO  
AND IN CHILDREN'S MAGAZINES AND COMIC BOOKS<sup>1</sup>

Jay D. Lindquist, Western Michigan University

Abstract

This is a preliminary report of exploratory research on the attitudes of elementary school children toward advertising on television and radio and in certain types of children's magazines and comic books. The results indicated that advertising in "quality" children's magazines was scored most positively while TV commercials were viewed most negatively by all grade levels tested. Analysis was carried out using appropriate ANOVA methods and t-tests.

Introduction

Published research on children's attitudes toward various media goes back to the 1930's when efforts were carried out to determine the effects of motion pictures on children. Next, attention was devoted to the criminal theme content of television programs in the mid-1950s and its impact on youngsters. Also since the 1950's there have been periodic investigations of television advertising and its effects on young people.

In 1950 the public at large was rather apathetic concerning the television-child relationship. Brumbaugh (1954) characterized the state of affairs as one where parents were not aware of the impact made upon a child's memory by advertising information concerning products in which a child had no immediate interest. Then (Thompson, 1964) it was found that television seemed to have little effect on the relationships of knowledge, understanding and acceptance of television commercials and actual use of the products advertised. It was then discovered (McNeal, 1964) that there was increasing dislike and mistrust of television advertisements as children grew older. In fact, one-half of the five and seven year olds tested and three-quarters of the nine year olds reported negative feelings toward television commercials. In 1970 James conducted a survey of children in grades four through twelve. He found that mean attitude scores were lower for television than any other medium and those scores lowered as a child's grade in school increased. He also concluded that as grade level in school increased the amount of time spent watching television decreased. This led to the inference that there was no positive relationship between attitudes toward television advertising and the amount of exposure to the medium. Then it was found (Ward, Reale and Levinson, 1971) that older children show a greater understanding as to what a commercial is and also can more easily discriminate between program material and commercials than younger children. Also research (Ward, Wackman and Levinson, 1971) showed a negative attitude development pattern toward TV commercials as children aged. Other work of interest has been done by Blatt, Spencer and Ward (1972); Robertson and Rossiter (1974); Ward, Wackman, Faber and Lesser (1974); Atkin (1975); Bever, Smith, Bengen, and Johnson (1975); Ferguson (1975) and Rossiter (1977).

<sup>1</sup>This research was supported by a fellowship and grant from the Faculty Research Fund, Western Michigan University.

Objectives

The objectives of the research are:

1. Measure the attitudes of third, fourth, fifth and sixth graders toward advertising on television and radio and in comic books and certain children's magazines.
2. Determine if differences exist within each grade across media.
3. Determine if differences exist within each medium across grade.

Methodology

Measuring Instrument

The seven-item, four-point scale instrument proposed and used by Rossiter (1977) was used to gather data on attitudes toward television commercials. This basic instrument, with minor modifications, was then used for the data acquisition on each of the other media. This questionnaire was selected because of its apparent success at the hands of Rossiter and to revalidate its reliability in measuring young people's attitudes toward television commercials and establish reliability in measuring attitudes toward the other three media in the study. The reliability data has not been processed and analyzed at this time and will be reported at a later date.

The seven items on the questionnaire (as labeled by Rossiter) are as follows:

<u>Item</u>	<u>Acronym</u>
1. Perceived truthfulness	TRUTH
2. Potential annoying qualities	ANNOY
3. Objectivity in describing advertised products	GOODONLY
4. Overall likeability	LIKE
5. Perceived persuasive power	PERSUADE
6. Believability of characters in the advertisement	BELIEVE
7. Trustworthiness as guides to product purchase	BESTBUY

The reader should note that items 1, 4, 6 and 7 are positively oriented whereas items 2, 3 and 5 are negatively oriented when respondents show degree of agreement. Recall that a four-point agreement scale was used. The four-positions were "agree very much, agree, disagree and disagree very much."

Attitude Scoring

Each item was scored on a four-point scale with the numbers running from "1" to "4" in the direction from "agree very much" to "disagree very much". To compute an overall attitude score, called TOTATT by Rossiter, a simple sum of the individual item scores was carried out. The only adjustment was in a reversal of scores for items 2, 3 and 5. The range of scores possible was 7 to 28, with the lower overall score being an indication of a "strong" positive attitude toward a particular medium and the higher score showing a "strong" negative attitude toward advertising in a particular medium.

Sampling Method

The sample was drawn from the Portage, Michigan portion of the Kalamazoo-Portage SMSA. The eleven elementary schools in the Portage School District were sorted into four groups based upon the predominant socio-economic classes within them. One school was then drawn at random from each group. Then one third, fourth, fifth and sixth grade class was selected at random from each of the four schools. This produced four grade cells with four schoolroom classes in each. The data used as the basis for this paper were collected from all four schools in the case of the television, radio and magazine media. However only two schools were used to gather comic book information.

The sample cells were:

Grade	Television-Radio-Magazines <sup>a</sup>	Comic Books <sup>b</sup>	Media <sup>c</sup>
Third	85	45	40
Fourth	88	38	33
Fifth	98	49	46
Sixth	94	47	43

n = 365                      n = 179                      n = 162

<sup>a</sup>Average pupils per cell. <sup>b</sup>Only two schools were given the chance to evaluate comic book advertising attitudes during the initial data gathering wave. However, on the retest, not reported here due to deadline requirements, all schools participated. <sup>c</sup>Respondents with acceptable questionnaires for all media.

Data Gathering Procedure

Data was gathered from the young respondents during their normal school day. Each class in each school filled out the instruments within their "home" room. The fourth, fifth and sixth graders participated on a self-administered basis with minor introductory remarks prior to each medium evaluation. The instruments were read on an item by item basis to the third graders.

Prior to beginning each of the questionnaires, the young people were asked to indicate whether they were "a boy or a girl" on the form. Before beginning the instrument concerning children's magazines, examples of the type to be considered were shown. These included Daisy, American Girl, Boys Life, National Geographic World and Ranger Ricks. Also the children were asked to indicate whether or not they had looked through or read any of them or any magazines like them. Covers of these magazines were mounted on a poster board and shown to the youngsters. Similarly a series of adventure and humorous comic book covers were mounted and shown to the school children prior to their completion of the comic book questionnaire. All respondents filled out the instruments on each medium in one session.

Findings

Item Analysis

Tables 1 through 4 are summaries of the item distributions for each of the four media. Within each table is a breakdown on the basis of grade in school. To facilitate a rough analysis by the reader, these tables have been collapsed into one "agreement-disagreement" table. The latter table is Table 5.

Examination of Table 5 reveals some interesting information. First note that advertising in children's magazines is seen as most truthful, followed in order by radio commercials, television commercials and advertis-

TABLE 1

ADVERTISING ON TELEVISION  
ITEM RESPONSE FREQUENCIES (%)

Item	Agree Very Much				Agree				Disagree				Disagree Very Much			
	Grade				Grade				Grade				Grade			
	3	4	5	6	3	4	5	6	3	4	5	6	3	4	5	6
TRUTH	12	3	3	2	37	42	32	22	38	39	54	53	13	16	11	22
ANNOY	27	23	6	13	22	26	31	36	22	36	50	48	29	16	13	3
GOODONLY	67	56	56	62	14	31	36	30	10	9	4	5	10	14	4	2
LIKE	16	15	4	3	32	32	54	48	24	30	32	33	28	24	10	16
PERSUADE	47	44	32	39	28	30	33	38	14	19	25	19	11	7	10	3
BELIEVE	14	2	5	3	16	10	9	8	27	49	51	45	43	39	35	44
BESTBUY	16	1	3	3	12	12	10	9	38	39	54	46	34	47	33	43

TABLE 2

ADVERTISING ON RADIO  
ITEM RESPONSE FREQUENCIES (%)

Item	Agree Very Much				Agree				Disagree				Disagree Very Much			
	Grade				Grade				Grade				Grade			
	3	4	5	6	3	4	5	6	3	4	5	6	3	4	5	6
TRUTH	21	10	4	3	35	27	42	37	32	55	48	49	12	8	6	11
ANNOY	28	10	7	3	22	23	24	20	20	48	56	59	29	19	13	18
GOODONLY	47	38	33	33	32	36	41	49	9	16	20	16	12	10	5	2
LIKE	37	21	17	21	21	42	43	47	19	22	26	24	24	16	13	8
PERSUADE	33	22	15	15	29	33	39	51	23	31	39	31	16	15	6	3
BELIEVE	12	6	1	3	17	17	17	12	35	43	55	61	36	34	27	25
BESTBUY	17	2	5	4	13	11	9	12	34	47	57	54	37	40	29	30

TABLE 3

ADVERTISING IN COMIC BOOKS  
ITEM RESPONSE FREQUENCIES (%)

Item	Agree Very Much				Agree				Disagree				Disagree Very Much			
	Grade				Grade				Grade				Grade			
	3	4	5	6	3	4	5	6	3	4	5	6	3	4	5	6
TRUTH	33	18	0	4	4	16	27	13	38	32	47	60	24	34	27	23
ANNOY	9	11	16	13	16	11	22	21	29	18	41	51	47	61	20	15
GOODONLY	29	26	33	28	24	26	51	49	16	18	10	17	31	29	6	6
LIKE	62	57	27	21	29	24	43	30	7	11	18	40	2	8	12	9
PERSUADE	11	13	35	36	29	18	31	36	27	26	27	15	33	42	8	13
BELIEVE	33	16	10	6	29	16	20	6	18	37	45	51	20	32	25	36
BESTBUY	27	11	6	2	11	5	12	2	27	42	51	43	36	42	31	53



TABLE 4

ADVERTISING IN CHILDREN'S MAGAZINES  
ITEM RESPONSE FREQUENCIES (%)

Item	Agree Very Much				Agree				Disagree				Disagree Very Much			
	Grade				Grade				Grade				Grade			
	3	4	5	6	3	4	5	6	3	4	5	6	3	4	5	6
TRUTH	35	25	5	6	41	48	69	45	16	23	20	39	10	5	6	10
ANNOY	12	1	2	3	6	7	14	16	26	45	58	55	56	47	26	26
GOODONLY	37	32	24	31	36	25	38	40	17	28	31	22	11	15	6	8
LIKE	63	38	21	15	25	47	52	52	10	12	18	28	2	3	9	5
PERSUADE	15	17	15	20	30	17	36	40	28	33	40	30	28	33	9	10
BELIEVE	28	9	5	3	31	35	31	12	33	31	48	57	8	25	16	28
BESTBUY	22	10	3	4	22	9	15	5	27	44	49	56	30	36	33	34

ing in comic books, if the differences among the percentages are significant. With the same disclaimer one may hypothesize that the table shows that advertisements in children's magazines are liked more often than those in comic books or on radio and television in descending order. Note also that these young people appear to feel that advertising in children's magazines does not encourage them to buy things that they don't really need to the degree that comic books, radio and television advertising do. Also of interest is the progressive feeling that products advertised in/on television, radio, comic books and children's magazines are the best products to buy. There is no intent on the part of the author to imply that the percentages shown are different in a statistically significant sense--a later analysis will treat this question. However, the data hints that such differences may exist.

TABLE 5

AGREEMENT/DISAGREEMENT SUMMARY  
BY MEDIA FOR ALL GRADES

Item	Agreement				Disagreement			
	Media				Media			
	TV	Radio	Books	Child Mag	TV	Radio	Books	Mag
TRUTH	38	45	28	68	63	55	72	32
ANNOY	46	34	30	15	54	66	70	85
GOODONLY	86	77	68	66	14	23	32	35
LIKE	51	62	72	78	49	38	28	22
PERSUADE	73	59	45	48	27	41	47	52
BELIEVE	17	21	34	38	83	79	66	62
BESTBUY	17	18	19	22	84	82	81	78

The agreement percentages are the sums of the "agree very much" and "agree" shares. The disagreement percentages are the sums of the "disagree very much" and "disagree" shares.

The agreement data presented in Table 6 is offered to point out certain possible trends that appear to be surfacing if one holds medium constant and looks across

grade. In most cases it may be hypothesized that for each of the four media, children's attitudes toward advertising within them become progressively more negative as grade in school (age) increases--not an unexpected finding. Additional analysis to more clearly determine the validity of such observations was carried out and will be presented later on in the paper.

TABLE 6

AGREEMENT ITEM RESPONSE FREQUENCIES  
AS A FUNCTION OF MEDIUM AND GRADE (%)

Item	TV Grade				Radio Grade				Comic Bks. Grade				Child Mags. Grade			
	3	4	5	6	3	4	5	6	3	4	5	6	3	4	5	6
	TRUTH	49	45	35	24	56	37	46	40	37	34	27	17	76	73	74
ANNOY	49	49	37	49	50	33	31	23	25	22	38	34	18	8	16	19
GOODONLY	81	77	92	92	79	74	74	82	53	52	84	77	73	57	62	71
LIKE	48	47	58	51	58	63	60	68	91	81	70	51	88	85	73	67
PERSUADE	75	74	65	77	62	55	54	66	40	31	66	72	45	34	51	60
BELIEVE	30	12	14	11	29	23	18	15	62	32	30	12	59	44	36	15
BESTBUY	28	13	13	12	30	13	14	16	38	16	18	4	44	19	18	9

These agreement percentages are the sums of the "agree very much" and "agree" shares.

Total Attitude Score (TOTATT) Analysis

In order to determine the positions of the various media in an overall positive to negative attitude sense, a look at average TOTATT by medium and by grade-in-school was taken. The data summarized in Table 7 was the basis for the next series of data presentations and analyses. Note that only the data collected from those students who scored all four media will be discussed from this point forward.

TABLE 7

MEAN TOTAL ATTITUDE SCORES (TOTATT)  
VERSUS GRADE IN SCHOOL AND MEDIUM

Grade		Children's			
		Television	Magazines	Radio	Comic Books
Third	n	40	40	40	40
	Mean	18.475	14.475	16.675	15.450
Fourth	n	33	33	33	33
	Mean	20.273	15.030	18.182	17.061
Fifth	n	46	46	46	46
	Mean	19.870	17.413	18.891	19.500
Sixth	n	43	43	43	43
	Mean	21.047	19.000	19.116	20.535

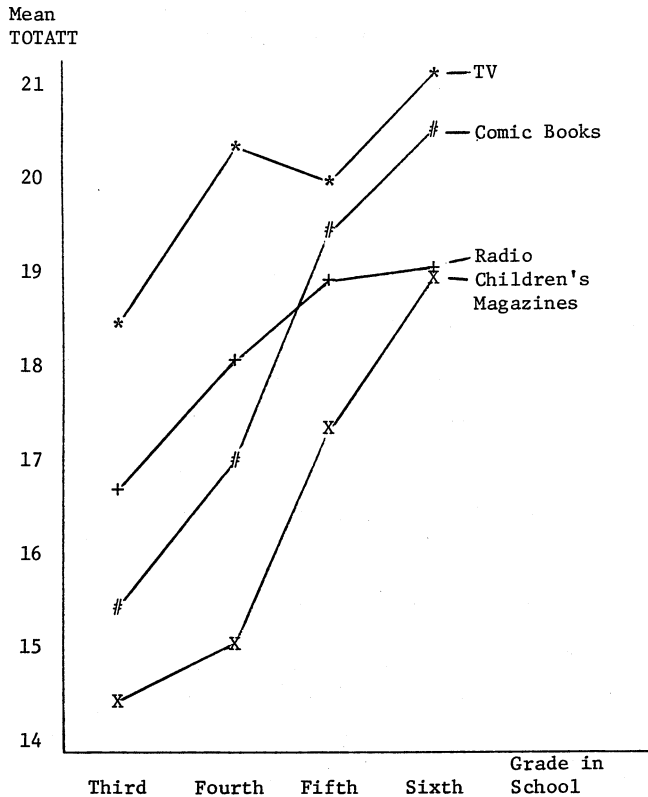
n = number of respondents who scored all four media within each grade.

A plot of the mean TOTATT scores for each medium across grades is presented in Chart 1. This preliminary look

at the data in descriptive form points toward nonlinear interaction within media as we proceed across grades. It could also be hypothesized from the chart that across grades the most favorable attitude is toward advertising in children's magazines (lowest mean TOTATT scores) and the least favorable is toward advertising on television.

CHART 1

TOTAL ATTITUDE SCORE MEANS  
VERSUS MEDIUM AND GRADE LEVEL



Also, hypothetically, we see the third and fourth graders favoring print media over sight and sound. One could also theorize that there is a shift in media position from fourth to fifth grade where advertising in comic books becomes more negatively viewed than radio commercials. Finally the chart allows us to speculate that the greatest shift in attitude is that associated with advertisements in comic books. Total attitude scores begin near the most positive of the four media and then by the time youngsters reach the sixth grade the comic books score is very near that of television, a much more negative position. In this case the comic book TOTATT mean score shifts 5.085 units as compared to shifts of 4.525, 2.575 and 2.441 units, respectively, for children's magazines, television and radio.

To further pursue this question a two-way repeated measure analysis of variance computer program was run on the data. Significant interactions between grade and medium were found at the 0.001 level. Table 8 displays the results of this run. Naturally the grade and medium treatment effects are also significant, in this case at the 0.001 level. Since interaction did exist, steps to clear up the within grade and within medium results had to be taken. One-way ANOVA was chosen to shed additional light on the subject. To begin with, a one-way repeated measure ANOVA was run holding grade constant and changing media. The rationale for using this procedure was that the data had been gathered from the same respondents across all four media, hence repeated measures existed. All possible pairs of

TABLE 8

TWO-WAY REPEATED MEASURE ANOVA RESULTS

Source	Degrees of Freedom	Sum of Squares	Mean Square	F-Ratio	Prob.
Between Subjects Grade	161	5058.49	31.42		
Subjects Within Groups	3	1241.96	413.99	17.14	<.001
Within Subjects Medium	158	3816.53	24.16		
Grade By Medium	486	5155.25	10.61		
Medium By Subjects	3	880.14	293.38	34.53	<.001
Within Groups	9	247.37	24.79	3.24	<.001
Total	647	10213.74			

groups within each of the four ANOVA runs were then tested to see if statistically significant differences existed. The methodology used in the latter case was the correlated t-test. A summary of the F-ratios for each grade in school is presented in Table 9.

TABLE 9

ONE-WAY REPEATED MEASURE ANOVA RESULTS BY GRADE  
ACROSS ALL MEDIA (BASED ON MEAN TOTATT SCORES)

Grade	F-Ratio	F-Probability
Third	9.909	<.001
Fourth	19.280	<.001
Fifth	7.515	<.001
Sixth	6.523	<.001

Then an overview of the media-pair significance levels is offered in Table 10. The data points toward third graders having similar attitudes toward advertising on the radio and that found in children's magazines and comic books. However, a closer tie between the two print media was shown by the t-tests. Further these young people have a significantly different attitude

TABLE 10

CORRELATED T-TEST PROBABILITY SUMMARY  
GRADE IN SCHOOL ACROSS MEDIA<sup>1</sup>

	TV vs Child Mags.	TV vs Radio	TV vs Comic Books	Child Mags. vs Radio	Child Mags. vs Comic Books	Radio vs Comic Books
Third	<.01	<.05	<.01	<.01	NS <sup>2</sup>	NS
Fourth	<.01	<.01	<.01	<.01	<.01	NS
Fifth	<.01	<.05	NS	<.02	<.01	NS
Sixth	<.01	<.01	NS	NS	<.05	<.02

<sup>1</sup>The probabilities shown are correct for a two-tailed test.

<sup>2</sup>Probabilities greater than .05 in value were considered not significant (NS).

toward advertisements in these three media than toward television commercials. As children move into the fourth grade, we note from Table 10 that the radio-comic book similarity still persists. Differences between television and radio commercials still exist as do those between advertising in comic books and children's magazines. (The reader may wish to cross-check Table 10 with Chart 1 throughout this discussion.) As the student enters the fifth grade one notes that the link between radio commercials and comic book advertising is strengthening. The gap also appears to be closing between advertising in comic books and television commercials. As the sixth graders are looked at, television commercials are perceived with an attitude similar to that expressed toward comic books. Also the relationship between radio and children's magazines appears to be quite strong. Overall, stability is noted in attitude positioning of television commercial scores versus the scores for children's magazine advertising.

At this point an examination across grades as medium was held constant was in order. Table 11 presents the results of a standard ANOVA for this circumstance. This procedure was appropriate since each data bit within each medium was gathered from a different respondent. The F-ratios were significant statistically in each of the media cases. The t-test summary for all possible grade pairs, holding medium constant is presented as Table 12.

TABLE 11

ANOVA RESULTS BY MEDIUM ACROSS ALL GRADES  
(BASED ON MEAN TOTATT SCORES)

Medium	F-Ratio	F-Probability
Television	4.721	.003
Children's Magazines	15.173	.001
Radio	4.090	.010
Comic Books	14.148	.001

TABLE 12

T-TEST PROBABILITY SUMMARY  
MEDIUM ACROSS GRADE IN SCHOOL<sup>1</sup>

Medium	Third vs Fourth	Third vs Fifth	Third vs Sixth	Fourth vs Fifth	Fourth vs Sixth	Fifth vs Sixth
	Fourth	Fifth	Sixth	Fifth	Sixth	Sixth
Television	.02	.05	.001	NS <sup>2</sup>	NS	NS
Children's Magazines	NS	.001	.001	.003	.001	.05
Radio	NS	.004	.002	NS	NS	NS
Comic Books	NS	.001	.001	.007	.001	NS

<sup>1</sup>The probabilities shown are correct for a two-tailed test.

<sup>2</sup>Probabilities greater than .05 were considered not significant (NS).

From this table one sees that fourth, fifth and sixth graders have similar attitudes toward television commercials. Only children in third and fourth grade appear to have closely related attitudes toward advertising in children's magazines. Radio commercials are viewed similarly by third and fourth graders whereas the fifth and sixth grade youngsters cluster independently with the fourth graders on the same medium. Table 13 gives a clearer picture of all of these grade and media relationships. Also note that third and fourth graders

show similar attitudes concerning advertising in comic books. The fifth and sixth graders also show this similarity in attitude though in a much more negative sense than that displayed with respect to the radio commercial. Refer to Chart 1 again to substantiate the latter in a graphical sense.

TABLE 13

RELATIONSHIPS ACROSS GRADE HOLDING MEDIUM CONSTANT

Medium	Grade			
	Third	Fourth	Fifth	Sixth
Television	X	X	X	X
Child Mags.	X	X	X	X
Radio	X	X	X	X
Comic Books	X	X	X	X

The solid lines indicate the statistically significant relationships that appear in the t-test data.

### Conclusions

First let us consider the limitations of the study. The instrument used was tested for its reliability by Rossiter (1977) in the gathering of data from fourth through sixth graders with middle class backgrounds from a large eastern city. The instrument was sufficiently reliable when used to measure attitudes toward television commercials. The author used the questionnaire with appropriate modifications on three additional media and one lower grade level and collected the data in a smaller midwestern city. Although reliability analysis is presently underway, it is not completed at this time hence questions pertaining to this factor are still present. Generalizability is also questionable since the sample was drawn from a midwestern city using a stratified random sampling method. In light of these qualifications, the study is considered exploratory and the following conclusions are seen as theoretical explanations for the results found.

Of the four media, children have the most negative attitude toward advertising in the television medium. This could be caused by such things as high exposure levels that lead to feelings of intrusion, prominent societal negative attitudes and parental negative reinforcement.

The most positive attitude across grade levels was that toward advertising in better quality children's magazines. These feelings could be caused by such things as the magazine being associated with reputable organizations (i.e. Boy Scouts, Girl Scouts, National Geographic, etc.) or because subscriptions are often purchased by parents or grandparents, giving tacit adult approval. Certainly quality of the advertising could also impact on the attitudes formed. Finally, these young people could also be feeling that there is sanctity in the printed word.

It is interesting to note the smaller change in attitude scores that radio commercials achieve across grades. Radio appears to be a more credible advertising medium for the fifth and sixth graders, particularly for the latter group where the mean TOTATT scores are very close to those for ads in quality children's magazines. This might be caused by the increased reliance of these groups on the medium as they enter the music-dominated preteen years.

Another interesting phenomenon is the wide comic book advertising attitude score change across grades. These combinations of picture-dominant, picture-word presentations start out near children's magazines in positive attitude placement. This might be because the third and fourth graders are in the early stages of their reading lives. Another factor might be the feeling of value for printed materials in general. The fifth and sixth graders, however, greatly increase their attitude scores (more negative) concerning advertising in this medium. Why? Possibly, they now have the ability to truly discriminate between magazines and comic books. Further, the latter are viewed more negatively by their peers and parents. Also, the type and quality of the advertisements themselves may have a bearing on the shifting scores.

It is interesting to note just how well even the third graders appear to discriminate among the media and also how, with experience, these elementary school children rearrange these media as they become more experienced. Finally, one sees that in all cases advertising media advertising TOTATT scores become more negative as these young people progress through school. This is not unlike many previous findings.

The data shows the potential power of quality children's magazines as an advertising medium across the grades studied. Also shown is the potential, already realized by many, that radio as an advertising medium appears to possess for reaching the preteener.

As is typical of exploratory work, more questions are raised than answered. The conclusions that have been presented are in the form of a series of speculations that point toward further research. Many doors that could be opened have been offered to the reader for his or her consideration.

#### References

- C. K. Atkin, The Effects of Television Advertising on Children: A Survey of Children's and Mother's Responses to Television Commercials (Department of Communications, Michigan State University, 1975)
- T.G. Bever, M.L. Smith, B. Bengen and T.G. Johnson, "Young Viewers' Troubling Response to TV Ads," Harvard Business Review, 53(November-December, 1975), 109-120.
- J. Blatt, L. Spencer and S. Ward, "A Cognitive Development Study of Children's Reactions to Television Advertising," in Television and Social Behavior, Vol. 4, Television in Day-to-Day Life: Patterns of Use, edited by E.A. Rubenstein, G.A. Comstock and J.P. Murray (Washington, D.C.:U.S. Dept. of HEW, 1971).
- F.N. Brumbaugh, "What effect Does TV Advertising Have on Children?," Educational Digest, 19(1954).
- C.P. Ferguson, Preadolescent Children's Attitudes Toward Television Commercials (Austin, TX: Bureau of Business Research, University of Texas at Austin, 1975).
- D.L. James, Youth, Media and Advertising (Austin, TX: Bureau of Business Research, University of Texas at Austin, 1971).
- J.U. McNeal, Children as Consumers (Austin, TX: Bureau of Business Research, University of Texas at Austin, 1964).
- T.S. Robertson and J.R. Rossiter, "Children and Commercial Persuasion: An Attribution Theory Analysis," Journal of Consumer Research, 1(June,1974), 13-20.
- J.R. Rossiter, "Reliability of a Short Test Measuring Children's Attitudes Toward TV Commercials," Journal of Consumer Research, 3(March, 1977), 179-184.
- G.W. Thompson, "Children's Acceptance of Television Advertising and the Relation of Televiewing to School Achievement," Journal of Educational Research, 58 (December, 1964).
- S. Ward, G. Reale and D. Levinson, "Children's Perceptions, Explanations and Judgments of Television Advertising: A Further Exploration," in Television and Social Behavior, Vol. 4, Television in Day-to-Day Life: Pattern of Use, edited by E.A. Rubenstein, G.A. Comstock and J.P. Murray (Washington, D.C.: U.S. Dept. of HEW, 1971).
- S. Ward, D. Levinson and D. Wackman, "Children's Attention to Television Advertising," in Television and Social Behavior, Vol. 4, Television in Day-to-Day Life: Pattern of Use, edited by E.A. Rubenstein, G.A. Comstock and J.P. Murray (Washington, D.C.: U.S. Dept. of HEW, 1971).
- S. Ward, D.B. Wackman and G.S. Lesser, Effects of Television Advertising on Consumer Socialization (Cambridge, MA: Marketing Science Institute, 1974).

Roger L. Jenkins, The University of Tennessee

## Abstract

Most researchers in the area of family decision-making have equated family decision-making with husband-wife decision-making and have excluded or ignored the role of children. Through the use of focus-group interviews with parents and by data collected from 105 husband-wife couples, this exploratory study focuses on: (1) the perceived role of the children in family decision-making in the areas of furniture decisions, automobiles, groceries, life insurance, savings, general family decisions, and in vacation decisions, and (2) the relationship between children's influence patterns and various demographic, socio-economic, personality, and attitudinal variables.

## Introduction

Family consumption behavior and the process of joint decision-making by husbands and wives has been of keen interest to marketing and advertising strategists as well as to sociologists and social psychologists (Cox, 1975; Cunningham and Green, 1974; Davis, 1976, 1975, 1974, 1972, 1970; Ferber and Lee, 1974; Hempel, 1974; Munsinger, Weber, and Hansen, 1975; Safilios-Rothschild, 1969). The last decade has witnessed a substantial change in the family role structure and family decision-making which is, in turn, reflected in the marketplace.

These role structure changes affect existing products, services and marketing practices. Marketers must remain abreast of the impact of these family role changes and their concomitant consumer purchasing patterns in order to modify marketing and advertising strategies accordingly. Just as a marketing researcher thinks of market segments composed of people with certain common demographic or attitudinal characteristics, he can also describe families which have certain decision-making characteristics for the purpose of developing advertising appeals and selecting media.

The influence of children in family decision-making has been a much too neglected subject of inquiry. Practically all research has centered upon the husband-wife dyad (Cunningham and Green, 1974; Davis, 1976, 1975, 1974, 1972, 1970; Ferber and Lee, 1974; Hempel, 1974; and Munsinger, Weber and Hansen, 1975) and ignored the role of children. Children, as would be expected, have an important role in the Western nuclear family. Adult consumer behavior is the direct antecedent of child consumer behavior. Many researchers have been concerned with children's responses to different kinds of television programming and advertising appeals; both Wells and LoSciuto (1966) and McNeal (1964) have studied patterns of child behavior within the store setting. But there appears to be a void in the literature of family decision-making in exploring children influence.

## Scope of the Study

This exploratory study investigates parents' perceptions of the role of children in various decision areas: major appliances, automobiles, furniture, groceries, savings, life insurance, vacations, and general household decisions. Do parents perceive children to have influence in such areas as major appliances or automobiles? Do husbands and wives differ in their perception of children's influence patterns? Why do children have more influence in some families than in others? These are relevant and

important questions which must be answered before a general theory or model of family decision-making can become a reality. Too, have previous researchers been justified in equating family decision-making with husband-wife decision-making in given product areas?

## Study Sample

Two focus-group interviews, each with five husband-wife teams, were held to establish the number, range, and commonality of decision making activities and subdecisions to be explored with the study sample of 105 married couples via a self-administered questionnaire for each spouse. The definition of "influence" varied from one person to another. Some perceived only the "active" dimension of the word while others perceived the word to encompass both the "active" and the "passive" dimensions.

A professional marketing research firm recruited ten married couples with children for the questionnaire pre-test and another 105 married couples for the final data collection. The research firm randomly called families until they filled the researcher's specified quota of having one-third of the couples having at least one child age five or younger; one-third with at least one child between the age of six and twelve years; and the remaining one-third with at least one teenager at home. Each participating couple was paid twenty-five dollars to complete the questionnaire. The couples were brought, in grouping of twenty couples, to a central location where they simultaneously and independently completed self-administered questionnaires.

Two sample limitations were recognized, but were judged to be within the acceptable "cost" limits. One limitation was representativeness of the couples as a sample of the Columbus, Ohio population. The families in the sample tended to have slightly higher income and educational levels than the general parent population. However, the sample utilized here is certainly more representative of the parent population than other samples in the family decision-making literature where students or convenience samples are usually the norm.

Another limitation of utilizing couples, some of whom had participated in previous projects conducted by the research firm, is that of bias in the responses of the couples. Respondents might be (1) trying to look good, (2) trying to appear as an expert or someone different from whom they really are, or (3) bored, annoyed, or fatigued from answering questionnaires at more frequent intervals than perhaps their neighbors. Since only twenty respondents indicated they had been recruited by the same research firm to take part in a similar survey, the possible bias would appear to be minimal.

Relative influence of husband, wife and each child was measured separately for each subdecision area. Each spouse was asked to respond on a constant sum scale by allocating 100 points among the family members. In order to minimize bias, the researcher collected the responses for each subdecision area before he introduced the next subdecision. As the explanation of subdecision areas and the definition or relative influence had to be consistent from one decision area to another and from one twenty-couple setting to another, standardized oral instructions were prepared and read at each session with respondents.

Additionally, both male and female directors, with equal authority, assisted in the questionnaire administration in order to minimize any sex bias that might be introduced otherwise.

### Results

Dependent upon the decision category and the specific subdecision area, parents' perceptions of children's influence varied greatly. The results are summarized in Table 1. The following generalizations can be made from Table 1:

- (1) Husbands and wives as groups are highly congruent in the perception of their own influence in decision-making. This is a rather unique finding when compared to other decision-making studies which have found a high amount of incongruence in spousal perception of their own influence patterns. The researcher postulates, in part, that this results from a tighter methodology for measuring spousal influence in this study as compared to others.
- (2) Automobiles, family savings, and life insurance decisions are ones which are perceived by both spouses to be more husband-dominant; furniture, grocery, and most general family decisions related to money management are perceived to be more wife-dominant; and major appliances and vacation decisions are perceived to be areas in which husbands and wives share decision-making more equally.
- (3) In general, husbands more than wives perceived their children to be more influential in family decision-making.
- (4) In most product categories, both spouses did not perceive children to exert a high amount of influence in decision-making. Children were perceived to exert most influence in vacation decisions and least influence in major appliance decisions.

From these results, it would appear that previous researchers who have excluded children from these product areas and have equated family decision-making with husband-wife decision making have been justified.

Given the fact that both spouses perceived children to exert a high amount of influence in vacation decision-making, the researcher chose this decision area to explore further children's influence patterns across families.

#### Explaining Children's Influence in Vacation Decision-Making

Why are children perceived to have more influence in some families more than in others? Are older children perceived to exert more influence? Is the stage in the family life cycle related to perceived children's influence? Are the spouses' ages, educational levels, and incomes correlated with perceived children's influence? Will children have more influence in families which are highly mobile? Will children be perceived as having more influence in families whose parents have "traditional" attitudes toward marital roles or when parents hold more "contemporary" views of marriage? Will a highly authoritarian parent perceive the children as having less influence? How are parents' levels of self-confidence related to children's influence patterns in vacation decision-making? These are the exploratory questions addressed here.

The husband-wife decision-making literature provided insight into which variables might be related to various

TABLE 1

### Average Relative Influence of Husband, Wife and Children In Decision-Making As Perceived By Parents

Decision Category	Percentage Influence Pattern					
	Husbands' Perception n=105	Wives' Perception n=105	Husbands' Perception n=105	Wives' Perception n=105	Husbands' Perception n=105	Wives' Perception n=105
<b>(1) MAJOR APPLIANCE DECISIONS</b> (refrigerator, washing machine, etc.)						
a. The decision when to buy	41.4	42.5	57.0	56.7	1.6	.8
b. The decision of how much to spend	51.0	52.4	48.2	47.5	.8	.1
c. The decision of where to buy	52.8	51.1	47.1	48.9	1.1	.0
d. The decision of which brand	43.7	40.3	55.3	59.5	1.0	.2
e. The decision of which model	36.7	35.9	62.7	63.4	.6	.7
<b>(2) AUTOMOBILE DECISIONS</b>						
a. The decision of which make to buy (Chev., Ford, etc.)	73.4	73.3	24.9	25.7	1.7	1.0
b. The decision of which model	65.5	65.5	28.8	28.9	5.7	4.6
c. The decision of what color	50.8	50.0	45.5	47.0	3.7	3.0
d. The decision of when to buy	68.0	69.0	31.0	30.7	1.0	.3
e. The decision of where to buy	78.8	73.7	20.9	26.2	.3	.1
f. The decision of how much to spend	68.3	68.5	31.1	31.4	.6	.1
g. The decision of what accessories to get	63.8	62.3	33.4	35.1	2.8	2.6
<b>(3) FURNITURE DECISIONS</b>						
a. The decision of when to buy	31.8	42.3	64.6	55.2	3.6	2.5
b. The decision of how much to spend	44.7	48.1	54.0	51.2	1.3	.7
c. The decision of where to buy	36.9	39.0	62.3	60.0	.8	.1
d. The decision of which piece of furniture to buy	29.7	34.9	67.6	61.4	2.7	2.7
e. The decision of which style, color, and fabric	26.9	28.1	69.6	67.3	3.5	4.6
<b>(4) GROCERIES</b>						
a. The decision when to shop	18.0	15.0	78.2	83.2	3.8	1.8
b. The decision how much to spend	26.9	22.9	70.5	75.5	2.6	1.6
c. The decision on which store	19.5	16.4	9.1	83.3	1.4	.3
d. The decision on which grocery products to buy	18.9	20.9	71.8	72.1	9.3	7.0
<b>(5) FAMILY SAVINGS</b>						
a. The decision when to save	58.6	59.1	39.4	39.7	2.0	1.2
b. The decision how much to save	58.5	60.4	39.2	38.5	2.3	1.1
c. The decision how to invest savings	65.6	69.0	33.6	30.2	.8	.8
<b>(6) LIFE INSURANCE</b>						
a. The decision on when to buy	67.4	72.7	29.0	24.7	3.6	2.6
b. The decision on amount to buy	69.1	73.1	28.6	25.2	2.3	1.7
c. The decision on from whom to pay	71.7	75.9	26.9	23.9	1.4	.2
d. The decision on type of policy	68.2	74.6	30.0	24.2	1.8	1.2
<b>(7) GENERAL DECISIONS</b>						
a. Who selects the family doctor	30.9	28.2	65.4	69.6	3.7	2.2
b. Who keeps track of money and bills	40.9	42.6	58.8	57.4	.3	2.0
c. The decision on what to do with extra money	46.5	47.4	49.8	49.4	3.7	3.2
d. The activities the family will participate in jointly	37.6	40.5	43.9	42.4	19.5	17.1
<b>VACATION DECISIONS</b>						
a. Total vacation decision	40.3	40.4	36.4	36.4	23.3	23.2
b. Information Collection	50.0	49.3	43.4	43.6	6.6	7.1
c. Whether to take kids	31.3	30.8	35.2	32.2	33.5	37.0
d. How long to stay	57.3	63.3	31.4	25.3	11.3	11.4
e. Actual date of vacation	53.1	55.8	22.6	18.8	24.1	25.4
f. Mode of transportation	51.8	53.0	41.7	43.8	6.5	3.2
g. Amount of money to spend	47.1	53.1	41.1	39.1	11.8	7.8
h. Kinds of activities	32.3	32.8	32.1	29.5	35.6	37.7
i. Selection of lodging	44.9	45.2	37.7	36.7	17.4	18.1
j. Selection of destination points	49.0	42.7	37.4	37.7	22.6	24.6

influence patterns. Will some of the same variables help explain perceived children's influence? The independent variables which were investigated to determine if they could individually and collectively explain perceived children influence patterns were:

- X<sub>1</sub> - Number of years couple has been married
- X<sub>2</sub> - Age of each spouse in years
- X<sub>3</sub> - Difference in ages of couple, one spouse relative to the other
- X<sub>4</sub> - Number of preschool children age five or younger at home in family
- X<sub>5</sub> - Number of children at home in family between ages of 6 and 12
- X<sub>6</sub> - Number of teenagers at home in family
- X<sub>7</sub> - Number of hours per week spent on work and work related tasks away from home
- X<sub>8</sub> - Number of hours per week spent on outside-the-home hobbies and interests
- X<sub>9</sub> - Number of formal years of education of each spouse
- X<sub>10</sub> - Difference in education levels of each spouse, one relative to other
- X<sub>11</sub> - Income level of each spouse
- X<sub>12</sub> - Difference in income levels of spouses, one relative to the other
- X<sub>13</sub> - Number of years lived in present city
- X<sub>14</sub> - Number of times couple has moved at least fifty miles from one home to another within the past five years
- X<sub>15</sub> - Score on "Attitude Toward Marital Roles" variable (Scale used to get total score developed by Catherine C. Arnott, Nov. 1972, Journal of Marriage and Family and by Cunningham and Green, Journal of Marketing Research, 1975.)
- X<sub>16</sub> - Husband-wife difference on "Attitude Toward Marital Roles" variable
- X<sub>17</sub> - Generalized score on authoritarianism variable (Score based on Adorno's Balanced F Scale and used by M. B. Smith, Journal of Personality, 1965.)
- X<sub>18</sub> - Husband-wife difference on generalized score on authoritarianism variable
- X<sub>19</sub> - Generalized score on self-confidence variable (Scale constructed by Day and Hamlin, Journal of Sociology, March, 1964, and used by Gerald Bell, Journal of Marketing Research, (February, 1967,))
- X<sub>20</sub> - Husband-wife difference on self-confidence variable

Using the husbands' data, regression models were developed to investigate the relationship of these independent variables, individually and collectively, to perceived children's influence in vacation decision-making. The procedure was duplicated for wives' data. The dependent variable in each regression model was parent's perceived influence of children in a given vacation subdecision as derived from the constant sum scale discussed earlier.

Since there are ten vacation subdecisions, the dependent variable produces ten regression equations for husbands' data and ten for wives' data. As the relationships discerned in the husbands' data were generally the same as found in the wives' data, only the husbands' regression models are discussed here.

The researcher is interested in determining whether the independent variables developed earlier are able to collectively explain variability in influence patterns and whether individually the variables are statistically related to those patterns. Table 2 shows the results of the regression runs for the total vacation decision and nine subdecisions.

The first focus is in determining the relationship between the individual predictors and the criterion variable. Using partial F values, the hypothesis that each of the twenty independent variables is not related to children's influence in decision-making is tested on husbands' data. Whether the hypothesis is accepted or rejected at the specified level of significance is stated in Table 2 for the total vacation decision and nine subdecisions.

For husbands, family life cycle variables are significantly correlated with children's influence in the following decisions: total vacation decision, whether to take children, how long to stay, kinds of activities, and selection of destination point(s). In the total vacation decision, Table 2 shows children's influence increasing the greater the number of teenagers and elementary school children in the family. Length of marriage is a statistically significant variable in four vacation decisions for husbands' perception of children's influence. In these decisions for each spouse, the longer the couple have been married, the greater the children's influence. This finding is in consonance with the early findings that older children are perceived to have greater influence.

In three vacation subdecisions--information collection, taking children, and amount of money to spend--Table 2 indicates older husbands allocate less influence to children. For husbands and wives the difference in age variable, more than absolute age, is related to influence patterns. In families where wives are older than husbands, children are perceived to have more influence.

Neither absolute income for each spouse nor spousal difference in income is significantly related to parents' perceptions of children's influence. It is not true for education. Table 2 indicates the higher the level of education for husbands, the less influence they perceive children to have, all other variables held constant. The hypothesis that husbands' perception of children's influence is independent of husbands' education is rejected in six of ten vacation decisions. The difference in education of husband and wife relative to husband is also a statistically significant variable in explaining children's influence in five vacation decisions for husbands. The greater the difference in education of the husband relative to the wife, the greater the influence of children.

The amount of time spent away from home is a significant explanatory variable for both spouses' perceptions of children's influence. The greater the number of hours husbands spend on work and work-related tasks, the greater the children's influence. Hours per week spent on outside-the-home hobbies and interests is not an important variable for husbands.

Only one of the two mobility measures is important in husbands' perceptions. The hypothesis that the number of years the couple has lived in present city is not related to children's influence is rejected for husbands' data in four vacation decisions. The sign of the standardized Beta coefficient indicates the longer the couple have lived in their present city, the lower the children's influence.

Each spouse's attitude toward marital roles and spousal difference on this variable are not highly correlated with perceived children's influence. The personality

TABLE 2

Husbands' Perception of Children's Influence in Vacation Decision-Making

VARIABLES	TOTAL VACATION DECISION (1)			INF. COLLECTION (2)			WHETHER TO TAKE CHILDREN (3)			HOW LONG TO STAY (4)			ACTUAL DATE (5)			
	BETA COEF.	PARTIAL F	SIG. LEVEL	BETA COEF.	PARTIAL F	SIG. LEVEL	BETA COEF.	PARTIAL F	SIG. LEVEL	BETA COEF.	PARTIAL F	SIG. LEVEL	BETA COEF.	PARTIAL F	SIG. LEVEL	
1 NO. OF CHILDREN AGE 0-5			N.S.			N.S.			N.S.	-.35	6.08	.025			N.S.	
2 NO. OF CHILDREN AGE 6-12	.27	7.04	.01			N.S.			N.S.						N.S.	
3 NO. OF CHILDREN AGE 13-19	.35	8.61	.01			N.S.			N.S.	.22	3.18	.10	.22	2.75	.10	
4 NO. OF YRS FORMAL EDUCATION	-.33	5.16	.01			N.S.	-.26	2.76	.10	-.22	2.85	.10			N.S.	
5 AGE			N.S.	-.30	2.95	.10	-.27	2.81	.10				N.S.		N.S.	
6 NO. OF YRS MARRIED TO PRESENT SPOUSE			N.S.	.44	3.65	.10	.39	3.44	.10				N.S.	.32	2.95	.10
7 INCOME			N.S.			N.S.			N.S.				N.S.		N.S.	
8 NO. OF TIMES MOVED DURING LAST 5 YRS			N.S.			N.S.			N.S.				N.S.		N.S.	
9 SCORE ON AUTHORITARIANISM			N.S.			N.S.	.31	2.97	.10	.57	10.68	.01	.27	2.77	.10	
10 SCORE ON ATTITUDE TOWARD MARITAL ROLES			N.S.			N.S.			N.S.	-.36	4.72	.05			N.S.	
11 SCORE ON SELF-CONFIDENCE	-.30	4.51	.05			N.S.	-.24	2.79	.10				N.S.		N.S.	
12 HOURS SPENT ON WORK RELATED TASKS	.20	3.46	.10			N.S.			N.S.	.27	6.02	.01			N.S.	
13 HOURS SPENT ON OUTSIDE INTERESTS			N.S.			N.S.			N.S.				N.S.		N.S.	
14 YRS LIVED IN PRESENT CITY			N.S.			N.S.			N.S.				N.S.	-.19	2.79	.10
15 AUTHORITARIANISM DIFFERENCE SCORE			N.S.			N.S.			N.S.				N.S.		N.S.	
16 ATTITUDE TOWARD MARITAL ROLES DIFFERENCE			N.S.			N.S.			N.S.				N.S.		N.S.	
17 SELF-CONFIDENCE DIFFERENCE SCORE	.32	5.46	.025			N.S.			N.S.				N.S.		N.S.	
18 RELATIVE DIFFERENCE IN AGE	-.21	4.24	.05			N.S.	-.17	2.92	.10				N.S.		N.S.	
19 RELATIVE DIFFERENCE IN EDUCATION	.23	3.84	.05			N.S.	.26	4.34	.10				N.S.	.21	2.91	.10
20 RELATIVE DIFFERENCE IN INCOME			N.S.			N.S.			N.S.				N.S.		N.S.	
21 R <sup>2</sup> (USING ALL VARIABLES)	32.74%			13.55%			25.77%			27.77%			22.06%			
22 ADJ R <sup>2</sup> (USING ALL VARIABLES)	17.71			-5.77			9.18			11.62			4.64			
23 OVERALL F AND SIGNIFICANCE	2.05 .025			.65 N.S.			1.56 .10			1.61 .10			1.18 N.S.			

VARIABLES	TRANSPORTATION MODE (6)			AMOUNT MONEY TO SPEND (7)			KINDS OF ACTIVITIES (8)			LODGING (9)			DESTINATION POINT(S) (10)			
	BETA COEF.	PARTIAL F	SIG. LEVEL	BETA COEF.	PARTIAL F	SIG. LEVEL	BETA COEF.	PARTIAL F	SIG. LEVEL	BETA COEF.	PARTIAL F	SIG. LEVEL	BETA COEF.	PARTIAL F	SIG. LEVEL	
1 NO. OF CHILDREN AGE 0-5			N.S.			N.S.			N.S.				N.S.	-.23	2.81	.10
2 NO. OF CHILDREN AGE 6-12			N.S.	.17	2.79	.10	.32	12.65	.01				N.S.	.27	7.02	.01
3 NO. OF CHILDREN AGE 13-19			N.S.			N.S.	.31	7.94	.01				N.S.			N.S.
4 NO. OF YRS FORMAL EDUCATION			N.S.	-.21	2.79	.10	-.25	3.36	.10				N.S.	-.32	4.55	.05
5 AGE			N.S.	-.27	2.82	.10			N.S.				N.S.		N.S.	
6 NO. OF YRS MARRIED TO PRESENT SPOUSE			N.S.	.51	5.98	.025			N.S.				N.S.		N.S.	
7 INCOME			N.S.			N.S.			N.S.				N.S.		N.S.	
8 NO. OF TIMES MOVED DURING LAST 5 YRS			N.S.			N.S.			N.S.				N.S.		N.S.	
9 SCORE ON AUTHORITARIANISM			N.S.	.25	2.83	.10	.25	2.77	.10				N.S.		N.S.	
10 SCORE ON ATTITUDE TOWARD MARITAL ROLES			N.S.			N.S.	-.25	2.95	.10				N.S.		N.S.	
11 SCORE ON SELF-CONFIDENCE	.24	2.77	.10	-.29	3.73	.10			N.S.				N.S.		N.S.	
12 HOURS SPENT ON WORK RELATED TASKS			N.S.	.32	8.76	.01			N.S.				N.S.		N.S.	
13 HOURS SPENT ON OUTSIDE INTERESTS			N.S.			N.S.			N.S.				N.S.		N.S.	
14 YRS LIVED IN PRESENT CITY			N.S.	-.17	2.76	.10	-.25	6.93	.01	-.19	2.94	.10			N.S.	
15 AUTHORITARIANISM DIFFERENCE SCORE			N.S.			N.S.			N.S.				N.S.		N.S.	
16 ATTITUDE TOWARD MARITAL ROLES DIFFERENCE			N.S.			N.S.			N.S.				N.S.		N.S.	
17 SELF-CONFIDENCE DIFFERENCE SCORE			N.S.	.27	3.59	.10	.23	3.44	.10				N.S.		N.S.	
18 RELATIVE DIFFERENCE IN AGE			N.S.			N.S.	-.26	7.83	.01				N.S.		N.S.	
19 RELATIVE DIFFERENCE IN EDUCATION			N.S.			N.S.	.20	3.69	.10				N.S.	.26	4.50	.05
20 RELATIVE DIFFERENCE IN INCOME			N.S.			N.S.			N.S.				N.S.		N.S.	
21 R <sup>2</sup> (USING ALL VARIABLES)	22.28%			29.13%			45.41%			16.34%			31.46%			
22 ADJ R <sup>2</sup> (USING ALL VARIABLES)	4.91%			13.26			33.20			-2.36			16.14			
23 OVERALL F AND SIGNIFICANCE	1.20 N.S.			1.73 .05			3.49 .01			.82 N.S.			1.93 .025			

variables were important in understanding husbands' perceptions of children's influence in decision-making. The absolute scores on authoritarianism and self-confidence are far superior to husband-wife difference measures on the same variables. Table 2 indicates the less authoritarian the husband, the more influence he perceives children to have. Similarly, the hypothesis that husbands' self-confidence is independent of his perception of children's influence is rejected in four vacation decisions. The direction of the relationship is such that the more self-confidence, the greater the influence allocated to children in vacation decision-making.

The research question now addressed is whether the combination of Variables 1-20 are able to explain husbands'

perceptions of children's influence in vacation decisions. The hypothesis tested is:

$$H_1: \text{Children's influence in the vacation decision is independent of the set of Variables } X_1 - X_{20}.$$

Table 2 indicates for husbands this hypothesis is rejected in six of the ten vacation decisions. Hypothesis 1 is rejected at the .01 level of significance for the kinds-of-activities vacation decision. This combination of variables is also statistically significant in explaining husbands' perception of children's influence in the total vacation decision, selecting destination(s), amount of money to spend, and whether to take children.



Adjusted  $R^2$ , or proportion of total variance in influence explained by these variables ranges, for wives, to a high of 23.05 percent for deciding on kinds of vacation activities. For husbands, these variables were able to explain a high of 33.20 percent in the kinds-of-activities decision.

#### Summary

Children were perceived to exert minimal influence in the following major decision categories: furniture, major appliances, automobiles, groceries, family savings, life insurance, and decisions regarding selection of family doctor, keeping track of bills and money, and spending extra income. However, both spouses perceived children to be highly influential in deciding on what activities the family will participate in jointly, especially vacation decisions. In fact, some specific vacation subdecisions could be called "children dominant!"

Regression models investigated individually are collectively the relationship between perceived children influence and various independent variables. The models revealed several interesting relationships. The greater number of elementary school children and teenagers in the family, the more influence allocated to children for decisions concerning kinds of activity destination, and amount of money to spend. The older the husband, the greater influence allocated to children for vacation decisions relating to information collection, whether to take children, and amount of money to spend. The greater the number of years married, the more influence husbands allocated to children in these subdecisions: information collection, whether to take children, actual date of vacation, and amount of money to spend. It can be concluded from the correlation of family life cycle variables with perceived children's influence that children in families which are in the later stages of the life cycle have more influence in vacation decision-making. Older children, especially teenagers, are perceived to have more influence.

Income is not associated with either spouse's perceived influence of children's input into decision-making. However, for husbands, education is related. The more educated the husband, the less influence he perceives children to have in the total vacation decision and the following subdecisions: whether to take children, how long to stay, amount of money to spend, kinds of vacation activities, and destination point(s). This might be explained by reference to the husband's profession and education required for it. In families where the husband is highly educated and holds a professional or comparable position, the responsibilities to such are usually great and the family's alternatives revolve to a large extent around them. In such instances, the husband is likely to have considerable ability to decipher information on vacation alternatives. Likewise, the decisions on how long to stay and how much money to spend may depend less upon the children's influence, and more on husband's position.

However, the more educated the husband relative to the wife, the greater influence children are perceived to have in the total vacation decision and the same subdecisions as before: whether to take children, actual date of vacation, kinds of vacation activities, and destination point(s). For husbands, the greater the amount of time spent away from home for work, the greater children's influence in the total vacation decision, how long to stay decision, and how much money to spend decision. One might infer from these findings that husbands who spend a great deal of time away from the children and home because of work, feel some self-actuated guilt and perceive children as having more influence because of it.

The less authoritarian the spouse, the more influence he allocated to children. Also, the more contemporary the attitudes toward marital roles held by a spouse, the more

influence he allocated to children. Husbands who possessed greater self-confidence perceived their children even more influential.

It can be concluded that in vacation decisions in which children were perceived to exert considerable influence, the overall regression equation was significant. For example, children were perceived to have the most influence, relative to other subdecisions, in deciding upon the kinds of vacation activities for the family to participate. This is the same subdecision in which adjusted  $R^2$  was highest--for husbands 23 percent and for wives 33 percent.

According to husbands' perceptions, increased children's influence in the total vacation decision is correlated with greater numbers of children ages six to twelve and thirteen to nineteen, less husband education, greater husband self-confidence, greater hours spent by husband on work, greater self-confidence of husband relative to wife, and younger in years relative to wife.

Hopefully, this exploratory investigation of children's influence in family decision-making will stimulate other research in this important area. The role of the child as an influencer of consumer behavior might also be investigated by studying the interaction between mother and child, between father and child, and across product purchase areas.

#### References

- Arnott, Catherine C., "Husbands' Attitude and Wives' Commitment to Employment," Journal of Marriage and the Family (November, 1972), 673-684.
- Bell, Gerald D., "Self-Confidence and Persuasion in Car Buying," Journal of Marketing Research, 4 (February, 1967), 46-52.
- Berey, Lewis and Richard W. Pollay, "The Influencing Role of the Child in Family Decision Making," Journal of Marketing Research, 5 (February, 1968), 70-72.
- Bernhardt, K. L., "Husband-Wife Influence in the Purchase Decision Process for Houses." Unpublished Ph.D. Dissertation, University of Michigan, 1974.
- Blood, Robert and Donald Wolfe, Husbands and Wives: The Dynamics of Married Living (New York: The Free Press, 1960).
- Braden, Joseph, "Husband-Wife Influence in Major Household Economic Decisions." Unpublished Doctoral Dissertation, Graduate School of Business, Indiana University, 1970.
- Burchinal, Lee G. and Ward W. Bauder, "Decision-Making and Role Patterns Among Iowa Farm and Nonfarm Families," Journal of Marriage of the Family (November, 1965).
- Cox, Eli P., "Family Purchase Decision Making and the Process of Adjustment," Journal of Marketing Research, 12 (May, 1975), 189-95.
- Cunningham, I. C. M. and R. T. Green, "Purchasing Roles in the United States Family, 1955 and 1973," Journal of Marketing, 38 (October, 1974), 61-64.
- Davis, Harry L., "Dimensions of Marital Roles in Consumer Decision Making," Journal of Marketing Research, 7 (May, 1970), 168-77.
- \_\_\_\_\_, "Measurement of Husband-Wife Influence in Consumer Purchase Decisions," Journal of Marketing Research, 8 (August, 1971), 305-12.

- \_\_\_\_\_, "Determinants of Marital Roles in a Consumer Purchase Decision," Working Paper 72-14. European Institute for Advanced Studies in Management, April, 1972.
- \_\_\_\_\_, and B. Rigauz, "Perception of Marital Roles in Decision Processes," Journal of Consumer Research, 1 (June, 1974), 51-62.
- \_\_\_\_\_, "Decision Making Within the Household," Paper Presented to RANN Program, National Science Foundation, April, 1975.
- \_\_\_\_\_, "Decision Making Within the Household," The Journal of Consumer Research: An Interdisciplinary Quarterly, 2 (March, 1976), 261-275.
- Draper, N. R. and H. Smith, Applied Regression Analysis (New York: John Wiley and Sons, 1966).
- Dunsing, Marilyn M. and Jeanne L. Hafstrom, "Methodological Considerations in Family Decision-Making," Association for Consumer Research Proceedings, January, 1975.
- Ferber, Robert, "Family Decision Making and Economic Behavior: A Review," In E. Sheldon (ed.), Family Economic Behavior (Philadelphia: J. B. Lippincott, 1973).
- \_\_\_\_\_, and L. C. Lee, "Husband-Wife Influence in Family Purchasing Behavior," Journal of Consumer Research, 1 (June, 1974), 43-50.
- \_\_\_\_\_, "Applications of Behavioral Theories to the Study of Family Marketing Behavior," Faculty Working Paper No. 125, University of Illinois, 1973.
- \_\_\_\_\_, "Some Unanswered Questions on Family Decision-Making," Association for Consumer Research Proceedings, January, 1975.
- Granbois, Donald H., 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.
- Green, R. T. and I. C. M. Cunningham, "Feminine Role Perception and Family Purchasing Decisions," Journal of Marketing Research, 12 (August, 1975), 325-32.
- Haley, Overholser and Associates, Inc., Purchase Influence: Measures of Husband/Wife Influence on Buying Decisions, 1975.
- Hallenbeck, P. N., "An Analysis of Power Dynamics in Marriage," Journal of Marriage and the Family (May, 1966), 200-203.
- Heer, D. M., "Husband and Wife Perceptions of Family Power Structure," Marriage and Family Living (May, 1962).
- Hampel, D. J., "Family Buying Decisions: A Cross-Cultural Perspective," Journal of Marketing Research, 11 (August, 1974), 295-302.
- Kenkel, W. F., "Decision-Making and the Life Cycle: Husband-Wife Interaction in Decision Making and Decision Choices," Journal of Social Psychology, 54 (August, 1961), 255-62.
- \_\_\_\_\_, "Observational Studies of Husband-Wife Interaction in Family Decision-Making," in M. B. Sussman (ed.), Sourcebook in Marriage and the Family (Boston, Mass.: Houghton-Mifflin, 1963).
- Komarovsky, Mirra, "Class Differences in Family Decision-Making on Expenditures," in Nelson N. Foote (ed.), Household Decision-Making (New York: University Press, 1961).
- McNeal, James U., Children as Consumers. Austin, Texas: Bureau of Business Research, University of Texas, 1964.
- Munsinger, G. M., J. Weber, and R. W. Hansen, "Joint Home Purchasing Decisions by Husbands and Wives," Journal of Consumer Research, 1 (March, 1975), 60-66.
- Olson, D. H., "The Measurement of Family Power by Self-Report and Behavioral Methods," Journal of Marriage and the Family, 31 (1969), 545-59.
- Safilios-Rothschild, Constantina, "Family Sociology or Wives' Family Sociology? A Cross-Cultural Examination of Decision-Making," Journal of Marriage and the Family (January, 1969).
- \_\_\_\_\_, "The Study of Family Power Structure: A Review 1960-69," Journal of Marriage and the Family (November, 1970).
- Sheth, J. N., "A Theory of Family Buying Decisions," in J. Sheth (ed.), Models of Buyer Behavior: Conceptual, Quantitative, and Empirical (New York: Harper and Row, 1974).
- Turk, J. L., and N. W. Bell, "Measuring Power in Families," Journal of Marriage and the Family, 34 (May, 1972), 215-22.
- Wheels, William D. and Leonard A. LoSciuto, "Direct Observations of Purchasing Behavior," Journal of Marketing Research, 3 (August, 1966), 227-33.
- Wilkening, E. A. and D. E. Morrison, "A Comparison of Husband and Wife Responses Concerning Who Makes Farm and Home Decisions," Journal of Marriage and the Family, 29 (1967), 703-11.
- Wilkes, R. E., "Husband-Wife Influence in Purchase Decisions--A Confirmation and Extension," Journal of Marketing Research, 7 (May, 1975), 224-7.

# CHILDREN AS INFORMATION SOURCES IN THE FAMILY DECISION TO EAT OUT<sup>1</sup>

James E. Nelson, Montana State University

## Abstract

Children's involvement in the nuclear family decisions to eat out is examined relative to parental involvement across six decision stages. Stages include problem recognition, providing information, deciding on restaurant type, deciding on a particular restaurant, deciding how much will be spent, and making the final decision. Results indicate children over five are as involved as parents in recognizing the problem, providing information, deciding on restaurant type, and deciding on a particular restaurant. For all families, parents appear to reserve the right to make the final decision and decide how much is spent.

## Introduction

Little academic marketing research has been concerned with the relative involvement of children as information sources in family decision making. Yet grounds for such an interest can at least be found in the large numbers of families with children and the numerous marketing strategies which focus on children.

This study concerns children as information sources in the family decision to eat out. As information sources, children can be influential in recognizing the problem, in providing information, and in making the final decision. Their relative influence, vis-à-vis adults in the decision to eat out, is central to this study. That is, are children significant information sources? Does their relative influence vary with demographic family characteristics? Finally, does the relative influence of children affect family decision criteria?

## Earlier Research

These research questions relate to four earlier studies. Szybillo and Sosanie (1977) examined 190 New York families' last decisions to have dinner at a fast food restaurant and to go on a one day family trip. Mothers indicated whether or not husbands, wives, and children were involved in purchase idea initiation, obtaining and providing decision alternatives, and making the final decision. Counts of families showed children involved at approximately 60-80% of all decision stages.

Berey and Pollay (1968) studied 48 mother-child pairs for children's influence on the mother's purchase of breakfast cereal. The mother's child centeredness rather than the child's assertiveness was significantly related to cereal brand purchases ( $r_s = -.27, \alpha < .05$ ). As with the Szybillo and Sosanie research, no demographic correlates of this influencing process were investigated.

Ward and Wackman (1972) studied parental yielding for 22 product purchase categories. Yielding was positively related to the child's age ( $r = .20, \alpha < .01$ ) but unrelated to family size and social class. Yielding was product specific, a finding confirmed by Mehotra and Torges (1977).

The present study extends this past research by examining the relative involvement of children across stages of the decision process. It replicates the Ward and Wackman (1972) and Mehotra and Torges (1977) research by examining demographic characteristics associated with children's involvement. It extends past research by examining children's involvement and family decision criteria.

## Method

Interviewers stopped consumers who appeared to be over age 18 at three shopping center locations in Bozeman and Billings, Montana. Consumers who "frequently ate pizza out" were asked if they would be willing to participate in a research study. Each of the 450 who agreed were given an address label and asked to record their address, to which a mail questionnaire would be sent. Each were promised a coupon good for a small pizza at an unnamed local restaurant if they returned a properly completed questionnaire.

Usable questionnaires were received from 290 consumers (64.4%). From this group, 84 respondents reported eating their last pizza out as a nuclear family. The main data analysis was conducted on respondents from this group who provided complete data sets (approximately 70, depending on variables involved in the analysis).

Consumers were asked to report the degree of involvement (not involved, slightly involved, moderately involved, and very involved) for themselves, spouses, and child or children across six stages of the decision process. For example, the first involvement question stated, "Thinking back to the last time you went out to eat, indicate how involved the people below were in deciding to eat out in the first place. For each person below circle the appropriate response." The other five questions relating to stages of the decision process asked for involvement in providing information in making the decision, in deciding what type of restaurant to go to, in deciding on a particular restaurant, in deciding how much money would be spent, and in making the final decision to eat out. Degree of involvement was coded 1 through 4 for not involved through very involved respectively.

Consumers were also asked to indicate what things are important in deciding "what pizza restaurant you eat at." Fifteen salient dimensions or decision criteria were identified; consumers rated each on a five-point "importance scale." Anchors used were "not important, little importance, important, moderately important, and extremely important" with responses coded 1 through 5 for increasing levels of importance.

Finally, respondents indicated their age, sex, marital status, age of their youngest child living at home, family size, occupation of the household head, their last grade of school completed, and family gross income in 1976. In comparing to nuclear family population figures, sample families were slightly younger, larger, better educated, and had higher incomes.

<sup>1</sup>The author expresses appreciation to Professor Robert D. Abbott, Montana State University, for his assistance in data analysis in this paper.

Results

Overall, children in this study were reported as slightly less involved across stages in the decision process than by Szybillo and Sosanie. Table 1 below presents the comparison.

TABLE 1  
CHILDREN INVOLVEMENT IN THE FAMILY DECISION TO EAT OUT BY DECISION STAGE

Children Involvement as a % of All Instances		
Decision Stage	Szybillo and Sosanie (n=190)	Present Study (n=84)
Problem Recognition	93	85
Provide Information	93	76
Final Decision	88	74
Restaurant Type	90	87
Particular Restaurant	91	82
Amount to Spend	21	22

Data in Table 1 generally indicate a widespread involvement of children across decision stages except for deciding how much will be spent.

The data, however, could be misleading. It may be that children are involved in the majority of decision stages but to only a small degree. Nothing in Table 1 makes any reference to the quality of children involvement, only to quantity.

Table 2 sheds some light on this issue, showing average involvement scores for all family members. Significance levels and sample sizes are also shown for the test of no difference between repeated measure means on each decision stage (Morrison, 1967). As Table 2 shows, this null hypothesis is rejected for each decision stage.

TABLE 2  
FAMILY MEMBER INVOLVEMENT IN THE DECISION TO EAT OUT, BY STAGE IN THE DECISION PROCESS

Mean Involvement For:					
Decision Stage	Me	Spouse	Children	n	Significance level* (α)
Problem Recognition	3.4	3.3	3.0	69	.037
Provide Information	3.4	3.3	2.7	68	.000
Final Decision	3.5	3.4	2.6	68	.000
Restaurant Type	3.5	3.3	3.1	70	.026
Particular Restaurant	3.4	3.3	2.9	67	.016
Amount to Spend	3.2	2.9	1.3	59	.000

\*The null hypothesis is rejected if  $F > F_{\alpha, 2, n-2+1}$ .

Contrasts or multiple comparisons between involvement scores in Table 2 for each decision stage were conducted with Hotelling's  $T^2$  statistic (Morrison, 1967) using a significance level of .01. Briefly, involvement scores between parents for all decision stages are not significantly different. Children involvement scores, however, are significantly lower than those for either parent in providing information, making a final decision, and deciding how much money will be spent.

Children involvement scores are not significantly different than those for either parent in recognizing the problem, deciding on restaurant type, and deciding on a particular restaurant.

Again the data in Table 2 have potential to mislead. Intuitively, and from Ward and Wackman (1972), one would expect the extent of involvement would vary with child age. Parents with children under six would seem less likely to use their children as information sources than parents of older children. Tables 3 and 4 provide details, showing mean involvement scores for families whose youngest child is age five or less, and six or over, respectively. Significance levels and sample sizes are also shown for the test of no difference between repeated measure means on each decision stage.

TABLE 3  
FAMILY MEMBER INVOLVEMENT IN THE DECISION TO EAT OUT BY STAGE IN THE DECISION PROCESS, CHILDREN AGE FIVE OR LESS

Mean Involvement For:					
Decision Stage	Me	Spouse	Children	n	Significance level* (α)
Problem Recognition	3.5	3.3	2.8	35	.005
Provide Information	3.6	3.3	2.5	37	.000
Final Decision	3.5	3.3	2.4	36	.000
Restaurant Type	3.5	3.2	2.9	39	.002
Particular Restaurant	3.3	3.2	2.7	38	.012
Amount to Spend	3.4	3.2	1.2	34	.000

\*The null hypothesis is rejected if  $F > F_{\alpha, 2, n-2+1}$ .

Table 3 shows more pronounced involvement differences between this group of younger children and their parents than the sample as a whole in Table 2. All children involvement sources are lower in magnitude while parental involvement scores are generally the same.

Contrasts between parental involvement scores showed no significant differences ( $\alpha < .01$ ) across the six decision stages. Children involvement scores, however, are significantly lower than scores for the responding parent (column 2 in Table 3) across all decision stages except for deciding on a particular restaurant. Children involvement scores are also significantly lower than spouse involvement scores in providing information, making the final decision, and deciding how much would be spent. In the other three stages, children/spouse involvement score differences are in the predicted direction but reach significance at approximately the .20 level.

The picture changes almost completely for families whose youngest child is age six or older. Table 4 shows the data.

TABLE 4

FAMILY MEMBER INVOLVEMENT IN THE DECISION TO EAT OUT  
BY STAGE IN THE DECISION PROCESS,  
CHILDREN AGE SIX AND OVER

Decision Stage	Mean Involvement For:				Significance level* ( $\alpha$ )
	Me	Spouse	Children	n	
Problem Recognition	3.4	3.4	3.3	34	.881
Provide Information	3.2	3.3	2.9	31	.228
Final Decision	3.5	3.5	2.8	32	.031
Restaurant Type Particular	3.4	3.5	3.4	31	.554
Restaurant	3.4	3.5	3.2	29	.580
Amount to Spend	3.0	2.6	1.4	25	.000

\*The null hypothesis is rejected if  $F > F_{\alpha, 2, n-2+1}$ .

Only in making a final decision and deciding how much money will be spent are mean involvement scores significantly different.

For these two decision stages, contrasts of mean differences between parental involvement scores are not significant. However, children's involvement scores are significantly lower than those for either parent in making the final decision ( $\alpha < .05$ ) and deciding how much would be spent ( $\alpha < .01$ ). That these older children are still significantly less involved than parents in making the final decision substantiates Berey and Pollay's "gatekeeper" concept (1968). Parents appear to reserve the right to overrule a child's consumption decision.

Rather than splitting the sample into two groups on the basis of children's age, it is possible to examine directly the relationship between age and children's relative involvement. Correlation coefficients between children's involvement scores across all six decision stages and all five socioeconomic characteristics are given in Table 5 below.

TABLE 5

CHILDREN'S INVOLVEMENT CORRELATED WITH SELECTED  
SOCIOECONOMIC CHARACTERISTICS BY DECISION STAGE\* (n=84)

Decision Stage	Age of Youngest Child	Family Size	Occupation	Education	Income
Problem Recognition	.28	.31	-.07	-.11	.13
Provide Information	.22	.20	-.15	-.04	.08
Final Decision	.34	.17	.03	-.04	.16
Restaurant Type Particular	.25	.23	.03	.01	.24
Restaurant	.26	.26	.05	.05	.22
Amount to Spend	.25	.17	.01	-.14	.05

\*Spearman rank order correlation, correlations greater than .19 are significant beyond  $\alpha = .05$ .

Children's involvement is positively related to age of the family's youngest child and, in general, to family size. The first finding is consistent with Ward and Wackman (1972); the second is not. Neither measure of social class, occupation or education, is significantly

correlated with children's relative involvement, again consistent with Ward and Wackman (1972).

The income/children involvement correlations are generally higher than those between social class and children involvement, although significant for only two stages in the decision process. This finding indicates that income may be more powerful than social class in predicting children's involvement in the decision.

Another analysis of children's involvement in family decision making is to examine only their involvement scores across decision stages. For example, children may be more involved in recognizing the problem but less involved in making the final decision. Table 6 compares children's involvement scores across all decision stages.

TABLE 6

CHILDREN'S INVOLVEMENT BY DECISION STAGE<sup>2</sup> (n=65)

Decision Stage	Mean Children Involvement
Problem Recognition	3.1
Provide Information	2.8
Final Decision	2.6
Restaurant Type	3.2
Particular Restaurant	3.0
Amount to Spend	1.4

The test of no difference between child involvement scores across decision stages is rejected at  $\alpha < .001$  for  $F(5,60)$  d.f.

Contrasts between involvement scores in Table 6 show significantly less child involvement in deciding how much would be spent than in deciding any of the other five decision stages ( $\alpha < .01$ ). Children are also significantly less involved in making the final decision than in recognizing the problem and deciding on restaurant type ( $\alpha < .05$ ).

Tests similar to that reported above were also conducted for parental involvement scores. Null hypotheses of no difference between involvement scores across decision stages for both parents were supported ( $\alpha < .05$ ).

A final research question deals with differences in the importance of selected decision criteria in choosing a pizza restaurant for high and low child involvement families. High child involvement families (n=38) were defined as those indicating that their children were either moderately or very involved in all decision aspects except deciding how much would be spent. Low child involvement families (n=12) were defined as those indicating no or slight involvement for these decision stages.

Mean importance scores on each decision criterion for the sample as a whole, high involvement families, and low involvement families are shown in Table 7.

<sup>2</sup>Means in Table 6 may be compared to those for children in Table 2; the small differences are explained by the slightly different samples for both tables. That is, data in Table 6 are based on those 65 families who provided a complete set of child involvement scores across all decision stages. Data in Table 2 reflect those families who provided complete involvement scores for both parents and child within each decision stage.

TABLE 7

## IMPORTANCE OF SELECTED DECISION CRITERIA IN CHOOSING A PIZZA RESTAURANT

Decision Criteria	Entire Sample (n=84)	High Involvement Families (n=38)	Low Involvement Families (n=12)
Cleanliness	4.8	4.8	4.5
Taste of the food	4.9	4.9	4.9
Variety of the menu	3.9	3.8	3.9
Service	4.4	4.3	4.3
Popularity with children	3.7	4.0	3.2
Convenient location	3.1	3.1	3.1
Prices	3.7	3.8	3.7
Familiarity with the restaurant	3.0	3.0	3.3
Identified manager on duty	2.0	2.1	1.9
Friendly employees	4.2	4.2	4.3
Comfortable atmosphere	4.3	4.2	4.6
Availability of wine or beer	2.3	2.2	2.8
Availability of entertainment	1.8	1.8	1.8
Availability of thick crust pizza	2.2	2.1	2.6
Availability of desserts	1.9	2.0	2.1

Because of the lack of published research on the issue and small sample sizes, no hypotheses were stated and tested. However, it appears that high and low involvement families may differ in their importance ratings for popularity with children, familiarity with the restaurant, comfortable atmosphere, availability of wine and beer, and thick crust pizza.

As might be expected from results earlier in this study, low child involvement families in Table 7 were significantly smaller and had younger children than high child involvement families.

#### Conclusions

The overwhelming conclusion from this research is that children are significant sources of information in the family decision to eat out. Beyond the age of five, children are generally as involved in this decision process as their parents except in making a final decision and deciding how much will be spent.

It is important to recognize that all children show in Table 2 near moderate involvement in recognizing the problem, providing information, deciding on restaurant type, and deciding on a particular restaurant. Thus, it would appear that children may be managerially significant information sources despite their statistically significant lower involvement scores. Lending credence to this conclusion are data from this study which show only about one family in six indicating no children involvement in the decision stages above.

There appear to be several differences between high and low child involvement families on several decision criteria. Larger sample sizes are needed to confirm this appearance.

#### Limitations

A limitation noted earlier in this paper should be repeated: earlier research has shown results to be product specific. Thus, caution should be exercised in extending results of this study to similar consumer decisions.

The sample suffers from its self-selection and non-random procedures, the last consideration being common

to all earlier reported research. Self selection is not a severe problem given that decision makers are more interested in users than the population as a whole. And given that memory problems of the consumer were likely not extreme in qualifying themselves as eating "a pizza out frequently." Not reported earlier, sample members also indicated on the questionnaire eating pizza out a total of 163 times or an average of 1.9 times in the two weeks prior to being approached as a prospective sample member.

Additional sample limitations concern geographic and population density considerations. Of the two, population density is probably the most important. Both Bozeman and Billings, Montana, are small urban areas with populations of approximately 25,000 and 75,000 respectively. An example of the distortion would seem to be the relatively low importance rating given location of the restaurant as a decision criterion. This finding can be largely explained by the relative ease of driving to any pizza restaurant in both cities. Research in larger urban areas would likely find location more important.

A final limitation concerns the nature of the data used in the involvement and decision criteria analyses. All such data are between ordinal and interval in terms of measurement. As such, the use of F ratios is debatable but not unreasonable (McNemar, 1969). The frequent caution that results must be interpreted with caution is appropriate.

#### Implications and Future Research Suggestions

Because of the high overall child involvement in the decision process, marketing strategies focusing on children are appropriate for this industry. However, the family market may be different from young singles, for example, on important dimensions. Care should be taken that strategies which appeal to families do not alienate other segments.

For the family market, it would appear that advertisements featuring more than one child over age five involved with parents in problem recognition, providing information, and in choosing a restaurant type or particular restaurant would be effective. Parents might be shown alone in advertisements making a final decision and deciding how much will be spent. In-store promo-

tions and products focusing on children are also consistent with this research.

Future research in different geographic areas in larger urban areas should be undertaken as replication. A reasonable workable methodology is demonstrated here and such studies should not be difficult to conduct. Larger samples of high child and low child involvement families are needed especially to examine differences in decision criteria for these two groups. Perhaps families may be segmented on this dimension.

Exploratory research is needed on potential alienation of other consumer segments from child or family centered marketing strategies. The concept of consumer-store image consistency hints at this alienation.

#### References

Lewis A. Berey and Richard W. Pollay, "The Influencing Role of the Child in Family Decision Making," Journal of Marketing Research, 5 (February 1968), 70-72.

Quinn McNemar, Psychological Statistics 4th ed. New York: John Wiley and Sons, 1969.

Sunil Mehotra and Sandra Torges, "Determinants of Children's Influence on Mothers' Buying Behavior," in William D. Perreault, Jr., ed., Advances in Consumer Research, Volume IV (Atlanta: Association for Consumer Research, 1977) 56-60.

Donald F. Morrison, Multivariate Statistical Methods, New York: McGraw-Hill Book Company, 1967.

George J. Szybillo and Arlene Sosanie, "Family Decision Making: Husband, Wife and Children" in William D. Perreault, Jr., ed., Advances in Consumer Research, Volume IV (Atlanta: Association for Consumer Research, 1977) 46-49.

Scott Ward and Daniel B. Wackman, "Children's Purchase Influence Attempts and Parental Yielding," Journal of Marketing Research, 9 (August, 1972) 316-19.

## CHILDREN'S CONSUMER RESEARCH: A CALL FOR RIGOR

John R. Rossiter, University of Pennsylvania

### Abstract

Children's consumer research is clearly in a growth phase headed toward maturity as an important subfield of consumer behavior. However, children's consumer research will attain mature status only if greater attention is given to theoretical contributions, practical value, measurement considerations, and generalizability.

### Introduction

Over 20% of the nation's consumers are children, and although most consumer behavior textbooks continue to neglect children's consumer behavior, more and more studies are appearing at conferences and in the journals. So it seems just a matter of time before this important subfield of consumer behavior becomes fully established. But it is just a matter of time? The point of this discussion paper is to suggest that it is not just more studies that are needed, but better studies.

The four preceding studies are part of the growth trend in children's consumer research. The studies are used to illustrate the need for better studies of a more rigorous nature. Researchers are urged to plan and execute their studies more rigorously by considering four standard criteria for good research:

1. Theoretical contributions. In order to advance our understanding of children's consumer behavior, studies should proceed from a theoretical background and attempt to contribute to this body of theory. Several theoretical approaches are possible and desirable. Major theoretical perspectives are reviewed in Robertson and Feldman (1975) and Ward, Wackman and Wartella (1977).
2. Practical value. Theoretical contributions have their own virtue. But there is a great need, as well, for children's research which has practical value. Policymakers in particular look to academic researchers to provide research findings of practical relevance to the difficult policy decisions they have to make. Marketers, too, seek practical studies but doubtless conduct their own studies to meet their needs.
3. Measurement considerations. One reason why our understanding of children's consumer behavior has not advanced as quickly as it should is due to widespread inattention to measurement considerations. Different measures with varying degrees of validity (usually no more than face validity) are used to represent supposedly identical constructs. Reliability is hardly ever examined (Rossiter, 1977) with a consequent lack of concern about the accuracy of findings. Internal rigor is thus a major problem in research to date.
4. Generalizability. If children's consumer research is to mature we have to give more attention to generalizability across studies, in the external sense. This applies not only to sample selection, but also to the sampling of content areas within studies. Family decision areas, children's products, and children's advertising should be sampled with just as much care as the children themselves.

These criteria are discussed in relation to the studies comprising this paper session.

### Toddler's Toy Satisfaction

Bjorklund and Bjorklund (1978) present an interesting study of toddlers' (age 1 to 2½) satisfaction with number and type of toys.

Theoretical contribution. This study represents a test of Piaget's substages of sensorimotor development. The test is a fairly reasonable one and it fails to support the substage notion. In an area where almost everyone who includes age as a variable in children's consumer research believes that the ubiquitous main effect for age supports Piaget's theory, this is a refreshing outcome. It should contribute to the search for other theories to account for these types of results.

Practical value. The Bjorklunds' study has some practical value for policymakers and marketers in that it implies that you do not have to portray a lot of toys, for example in a commercial, in order to attract or satisfy a toddler. It would be interesting to extend the study to older, less egocentric children to see whether the common belief that older, more peer-oriented children are admired or envied for the number of toys they have holds true. Parents may dispute the practicality of the "satisfied with a few toys" finding, however, as they typically have to keep toddlers amused for longer than the 10-minute periods used in the study.

Measurement considerations. The use of behavioral measures of children's satisfaction shows appropriate recognition of the validity problem with prearticulate children, where verbal measures cannot be used. Even in studies with older, more articulate children there is a need for greater use of behavioral measures, especially to supplement or cross-validate verbal reports and verbally based questionnaire ratings.

Reliability is also acknowledged in this study. However, the use of coders who are familiar with the study's objectives reduces the faith that can be placed in the reliability estimate. But all too often principal investigators code their own data, a bad practice to begin with but unforgivable without inter-coder (parallel forms) reliability checks. At least an attempt was made to avoid this problem here.

Generalizability. Given the small sample size, particularly for each cell of the design, it is amazing that the ANOVA produced treatment effects which emerged through the subject variance. The study could be criticized, however, for using 12 to 20-month-old children when the population of interest comprises the 12 to 30-month-old age group. Also, the very strong upward SES bias limits generalizability and this should have been discussed; for example, in terms of I.Q. and attention span for toy play activities.

### Children's Media Attitudes

Lindquist (1978) gratifyingly follows a suggestion made by Rossiter (1977) that children's attitudes toward



television advertising should be compared with their attitudes toward advertising in other media, in this case radio, children's magazines and comic books. Lindquist's study improves on an earlier comparison by James (1971) by using a multi-item attitude measure with demonstrated reliability.

Theoretical contribution. This study provides a good example of how not to link theory with empirical data. Instead of thinking about the nature of each medium and developing some hypotheses about how children might perceive advertising in the respective media, the author offers the hypotheses after the data have been presented! Also, no attempt is made to relate the findings to those of James (1971) or Rossiter (1977) and thereby to consolidate knowledge in this area.

Practical value. Children's attitudes toward various media as advertising vehicles may be of some use to advertisers (though I doubt it based on the validity issue discussed below). On the other hand, Lindquist's study does have a certain relevance for public policy, if only to remind policymakers that television commercials are not the only source of advertising to which children are exposed.

Measurement considerations. The 7-item attitude measures used by Lindquist were either identical to (in the case of TV advertising) or directly adapted from (in the case of the other three media) the measure developed by Rossiter (1977). The measure was shown to have acceptable internal consistency reliability and test-retest reliability, the two appropriate types of reliability in this instance, for attitudes toward TV advertising measured on a sample of 4th through 6th grade children. Unfortunately, Lindquist did not have time to compute comparable reliability estimates for the present data. I, for one, anxiously await these computations. Others should too, for this is the only self-administered attitude measure in the children's consumer research literature for which reliability estimates are available. The sooner we begin to use standardized measures of known reliability, the better our results are going to get, simply because of reduced error variance.

The crucial issue, however, concerns the validity of children's attitude measures. While there is little doubt that the attitude measure used in this study taps the construct of attitude adequately (construct validity) and samples the traditional cognitive, affective and motivational components of attitude (content validity), there is considerable doubt as to its predictive validity. There is no evidence that increasingly negative attitudes toward television commercials as children grow older bear any relationship to actual behavior, such as requests for advertised products (Rossiter, in press). It is unlikely that the picture would be any different for other media. Children simply seem to learn more cynical verbally expressed attitudes but these attitudes do not seem to affect their behavior. Advertisers, therefore, would be ill advised to place stock in these types of findings.

Generalizability. As noted, it would have been interesting to see to what extent Rossiter's (1977) results generalized to a somewhat broader sample (3rd through 6th grade versus 4th through 6th, and a wider SES spread) and to other media. Also, given the Federal Trade Commission's focus on age 7 as a cutoff for "total" protection from TV advertising--a focus based, incidentally, largely on attitude data--extension of the study to 1st and 2nd grades would be useful. This can be accomplished by using a reader to assist younger children, as Lindquist did at 3rd grade, or by administering the test as a series of open-ended questions and using multiple coders, as recommended in Rossiter (1977).

The studies by Jenkins (1978) and Nelson (1978) examine the role of children in family decision making. As such, they can be considered together.

Theoretical contribution. Both studies are essentially atheoretical. This is a pity because a little forethought as to what children's "influence" (Jenkins) or "involvement" (Nelson) in family decision making entails might have prevented the largely uninterpretable findings that resulted in both instances. Surely any study of children's roles in family decision making must acknowledge the difference between what Haley, Overholser and Associates (1975) termed direct influence and indirect influence. The former represents an active role based directly on the decision maker's own needs whereas the latter represents a passive role in which the decision maker takes another family member's needs indirectly into account. This distinction was developed for husband-wife dyads, but it would seem to be crucial when children are included in the decision process (cf. Well's 1965 concept of "passive dictation"). In short, neither study contributes theoretically to our understanding of the role (or roles) of children in family decision making.

Practical value. As straight empirical studies Jenkins' and Nelson's efforts may have had some practical value to marketers were it not for definitional ambiguities, particularly in Jenkins' case, and sampling problems, particularly in Nelson's case. These problems are discussed below.

We should not, however, lose sight of the importance of more research on family decision making in the true sense--that is, including children, when appropriate. As Jenkins and Nelson note, far too many studies of family decision making focus exclusively on adults.

Measurement considerations. It is impossible to obtain valid measures of constructs such as "influence" or "involvement" unless these constructs are clearly defined at the outset. Jenkins freely admits in his pretest that interpretation of the word "influence" varied from one person to another, yet the study proceeded with this fatal ambiguity unresolved. Nelson apparently decided to avoid this complication for his concept of "involvement" by not pretesting at all.

Unreliability is likely to have entered the measures (quite apart from the basic validity problem) in a number of ways. The data in both studies rely on parental reports; and while it has been shown that, on an aggregate basis, husbands can fairly reliably report for wives and vice versa, this has by no means been demonstrated for parents reporting for children (Rossiter and Robertson, 1975). Other problems include the use of single item measures and reliance on recall, the latter severe in Jenkins' study but not so much in Nelson's, although Nelson is at fault for trying to use monadic measures rather than ratio measures, essential for the notion of joint decisions. Finally, both studies introduce "noise" by including teenagers as potential role-takers when they purport to be studies of children.

Generalizability. Earlier in this discussion it was remarked that content generalizability is just as important as sample generalizability. Each study provides an illustration.

Jenkins selected eight content areas to represent family decision making: major appliances, automobiles, furniture, groceries, savings, life insurance, vacation decisions and a miscellaneous category. How and why were these selected? Only one content area (vacations)

turned out to have substantial children's influence. This is not too surprising a result unless one were to have distinguished direct influence from indirect influence. It is hard to imagine parents deciding on automobiles, furniture, or life insurance, for example, without taking children indirectly into account. Neglected are content areas where children a priori are known to have a greater role in purchase decisions (Ward and Wackman, 1972). Without adequate sampling of content we cannot hope to make meaningful generalizations about family decision making.

Nelson's study contains the more traditionally recognized sampling problem: subject selection. Participation in his study was confined not only to people who frequently ate pizza out but also to those who reported eating their last pizza out as a nuclear family. Let me suggest that there may be a conspiracy among children to veto pizza dining in favor of McDonald's. The sampling problem should be obvious.

#### Conclusions

Children's consumer research has reached a stage of development in which increased rigor in planning and execution is of paramount importance. In the rush to fill the knowledge void concerning children's consumer behavior we have all been guilty of lack of rigor in particular areas. Four such areas are suggested for increased attention in future research:

1. Theoretical contributions. Lack of theory usually stems from lack of thought. Studies should be planned as contributions to an overall body of knowledge rather than as singular empirical ventures.
2. Practical value. At the present time the overriding need is for studies that have practical relevance for policymakers. The alternative is policies based largely on speculation accompanied by the allegation that we, the experts, have failed to deliver.
3. Measurement considerations. Validity and reliability are the hallmarks of sound research and accepted principles of scientific measurement cannot continue to be neglected, either in adult consumer research (Jacoby, 1978) or in children's.
4. Generalizability. Generalizability is the crux of science (see, for example, Nunnally, 1978). Sampling of content as well as sampling of subjects must be given greater attention in the interest of generalizable research results.

If children's consumer research should pass through a fad cycle, as so many of our research interests do, let us be sure that we leave a substantial and sophisticated contribution. Only by constructive efforts along the lines suggested above, at the inevitable price of a little destructive activity among colleagues, can we achieve this.

#### References

- G. Bjorklund and R. Bjorklund, "An Exploratory Study of Toddlers' Satisfaction With Their Toy Environments," paper presented at the Ninth Annual Conference of the Association for Consumer Research (Miami, Florida, 1978).
- R.I. Haley, C.E. Overholser and Associates, Purchase Influence: Measures of Husband/Wife Influence on Buying Decisions. Summary Report (New Canaan, Connecticut: Haley, Overholser and Associates, 1975).

J. Jacoby, "Consumer Research: A State of the Art Review," Journal of Marketing, 42 (April 1978), 87-96.

D.L. James, Youth, Media and Advertising (Austin, Texas: Bureau of Business Research, University of Texas at Austin, 1971).

R.L. Jenkins, "The Influence of Children in Family Decision Making: Parents' Perceptions," paper presented at the Ninth Annual Conference of the Association for Consumer Research (Miami, Florida, 1978).

J.D. Lindquist, "Children's Attitudes Toward Advertising on Television and Radio and in Children's Magazines and Comic Books," paper presented at the Ninth Annual Conference of the Association for Consumer Research (Miami, Florida, 1978).

J.E. Nelson, "Children as Information Sources in the Family Decision to Eat Out," paper presented at the Ninth Annual Conference of the Association for Consumer Research (Miami, Florida, 1978).

J.C. Nunnally, Psychometric Theory (New York, New York: McGraw-Hill, 2nd edition, 1978).

T.S. Robertson and S. Feldman, "Children as Consumers: The Need for Multitheoretical Perspectives," in B.B. Anderson (Ed.), Advances in Consumer Research: Volume III (Cincinnati, Ohio: Association for Consumer Research, 1975), 508-512.

J.R. Rossiter, "Reliability of a Short Test Measuring Children's Attitudes Toward TV Commercials," Journal of Consumer Research, 3 (March 1977), 179-184.

J.R. Rossiter, "Does TV Advertising Affect Children?," Journal of Advertising Research, (in press).

J.R. Rossiter and T.S. Robertson, "Children's Television Viewing: An Examination of Parent-Child Consensus," Sociometry, 39 (September 1975), 308-326.

G.J. Szybillo and A. Sosanie, "Family Decision Making: Husband, Wife and Children," in W.D. Perrault, Jr. (Ed.), Advances in Consumer Research: Volume IV (Atlanta, Georgia: Association for Consumer Research, 1977), 46-49.

S. Ward and D.B. Wackman, "Children's Purchase Influence Attempts and Parental Yielding," Journal of Marketing Research, 9 (August 1972), 316-319.

S. Ward, D.B. Wackman and E. Wartella, How Children Learn to Buy (Beverly Hills, California: Sage: 1977).

W.D. Wells, "Communicating with Children," Journal of Advertising Research, 5 (June 1965), 2-14.

RESEARCH ON MARKETING AND CHILDREN: UPSIDE OR DOWNSIDE ON  
THE PRODUCT LIFE CYCLE?

Scott Ward, Harvard University and Marketing Science Institute

Abstract

Session papers are reviewed in terms of three criteria which might be applied to indicate whether research on marketing and children will form an important and lasting sub-area of consumer behavior research, or whether research in this area is merely a temporary "hot topic."

Introduction

Children are a minority group around the world, and consumer researchers who study children are also a minority group. As such, they must endure certain indignities--witness, for example, that the only pejorative session title at this whole conference is applied to this session! Instead of dignified-sounding labels like "multi-attribute...", "multidimensional..." or "information overload," we are branded with, "from tots to tarts!" But the individuals who prepared papers for this conference, and the individuals who showed up at the session presumably have some common interests in what they see as a worthwhile set of consumer research problems. Is it worth the time and effort?

Recently, Alan Andreasen examined some historical trends in consumer research and he demonstrated that areas of research interest have a life cycle like products and people (Andreasen, 1978). The frequency of articles and papers dealing with marketing to minorities, for example, followed a sort of product life cycle curve over a several-year period. One thinks about other "seasons" of consumer behavior research topics: personality and buying behavior, diffusion of innovations, and so on. I think we'd pretty much agree that consumer research that somehow involves children is still in the "growth" stage of the life cycle; the question is, will studies in this area lead to a more or less permanent place in the systematic study of consumer behavior (and be freed of cute labels for convention sessions)?\* Is the evolving model closer to child psychology--a more or less permanent and important aspect of the study of human psychology--or is the model closer to narrow areas of study which are important, but temporary, since their only source of demand derives from current environmental issues and conditions?

I submit that the answer to this question depends upon the answers to three more basic questions:

1. Does the study of marketing and children give the field of consumer research a "new look?" That is, does it advance our knowledge in areas which the producers and users of consumer research find important and useful?
2. Do studies in this general area exhibit--and contribute to--conceptual clarity, and the development of theory?
3. Is there attention to relevance in research, to marketing practice, and/or to public policy decisions?

Let me take these general questions one at a time, and evaluate the papers at this session as specific object lessons.

\*[ED. NOTE: Professor Ward is author of the famous "KID'S TV -- Marketers on Hot Seat" article, Harvard Business Review, July 1972.]

1. Does Research Bring a "New Look?"

First, studies of marketing as it relates to children and families clearly have the potential for giving the field a "new look," and one which would seem to be potentially important and useful. The focus is on inter-individual processes--a trend away from earlier traditions in consumer research which focused exclusively on intra-individual processes while ignoring social processes. There is also great attraction in studying consumption behavior as it "actually occurs." We are not bound to contrived experiments with college sophomores, or to yet another tedious questionnaire mailed to housepersons. There is intuitive, as well as empirical support for the notion that many purchases are influenced by, or because of, "family" considerations, and children themselves are consumers whose attraction to marketers increases as they age. True, the birth rate is down, and singles and oldsters are emerging as other consumer phenomena worth studying, but the family seems to be a venerable social institution, with important economic consequences.

The authors of papers presented here recognize that previous research on family consumer behavior has largely focused on husbands and wives and ignored children. To the extent that research contributes to conceptual models which include offspring, we will have a "new look," indeed.

2. Conceptual Clarity and Theory Development

What about the second question: research in this area exhibiting conceptual clarity, and usefully applying, and contributing to, theory? We are blessed with a surfeit of theoretical approaches to the study of children which can help us understand their consumption behavior, and their role in family consumption behavior. Zigler and Child (1969) in their seminal review of socialization research, identify no less than seven major theoretical approaches to study in this area. Add to that various theoretical notions in areas designated something like "family sociology," and you have even more. The dominant theory in child psychology these days is cognitive development theory, so it is not surprising that this theory dominates research pertaining to marketing and children. It's useful in describing and explaining the thoughts and behaviors of children at different stages of the early life cycle, and the neo-Piagetians have usefully extended the theory to suggest some implications for consumer researchers who are by and large more interested in applications, e.g., the possibilities of training different-staged children to more fully understand advertising messages. Consumer research can not only "borrow" from existing social science theories, but usefully extend these theories--if we choose to do so.

A first step toward useful applications and useful extensions of theory is carefully thinking through concepts and their operational measures we employ. What do we mean, for example, when we speak about children's influence on family decision-making, as Jenkins and Nelson have done? Are we talking about (and measuring) relatively direct efforts of children to influence specific purchases? Or are we talking about the oft-cited phenomenon of "passive dictation," that Wells has discussed (Wells, 1965) in which some parental purchases are partly

predicted by their perceptions of children's unstated preferences? Are we talking about parental decision-making which occurs *de facto* because children are present, but in the absence of any "influence" from them (e.g., baby cribs and term insurance), or decisions adults make as parents although their children may not even be able to vocalize preferences before the fact (e.g., pre-schoolers being dragged around the Smithsonian)? If there is significant inter-subject variation in the definition of "influence" of children as Jenkins found, then it seems to me that the researcher ought to narrow and standardize that definition, so we know what results mean. That task, in turn, depends upon one's underlying, conceptual understanding of that thorny concept, "influence." The same can be said, incidentally, about the concept of "involvement," as in Nelson's paper. It's necessary to ask, influence on what, or involvement *in* what. Jenkins is interested in children's influence on a range of general and specific product and service categories that adults buy. But why were the decision categories chosen? They were probably used in previous research on husband-wife decision making, but it is difficult to understand why children would be expected to have much "influence" on major appliance, automobile, furniture, and savings decisions, especially if the concept of influence implied to some parents a notion of a relatively specific and direct request by children. Parent's decisions to buy furniture, insurance, and the like may be affected by the presence of children, and "influenced" in this way, but it's difficult to envision a kid asking Dad for a Maytag dryer or a Sealy Posturepedic.

Some conceptual and measurement issues also attend Nelson's study of children as information sources in the family decision to eat out. It is not clear how involvement is conceptualized, and how valid and reliable the measurement of the variable is. Depending on how respondents were defining the term, the following scenarios might each result in parents saying their kid was quite important, or highly involved, even though different processes--with different implications--were going on. Consider: the family is driving along and the kid is throwing a fit in the back seat because he/she wants to go to McDonald's. But a second scenario is: the family is driving along and one parent decides to stop because he/she intuites that his/her kid likes McDonald's and would enjoy having a hamburger. A third scenario might be: the family is driving along and even though the parents would vastly prefer to stop off at the Côte Basque, they stop at McDonald's because that's where they have the time and money to feed their brood.

Are these trivial distinctions? I think not. If we are to have a satisfactory conceptualization of intra-family influence processes, then the nature of that influence must be clearly defined and delimited. From a policy perspective, critics of marketing to children are always asserting that children have enormous influence on parental purchasing, and they are envisioning the back-seat tantrum variety. Other less malignant processes may be going on. On the other hand, it is misleading to conclude that the husband-wife researchers have been right all along, and children have little influence on purchasing, if we ask about purchases for which children should not be expected to have much influence, or about purchases which are defined so broadly (grocery products) as to be of little use.

There are some nitty-gritty issues in the Nelson and Jenkins studies. Briefly, in Nelson's study, data are based on retrospective reports from one family member's perspective; sample procedure was not random; the study did not control for which family member answered the questionnaire, and, as Jenkins' study shows, this can make a difference. Finally, I am not sure how valid and reliable the six decision stages are, despite their

popularity in other research. There would seem to be a lot of situational variables involved, which would lead to a lot of variation in how respondents interpreted their decision stages. In Jenkins' study, my main criticism--aside from my concern as to why the particular decision categories were chosen--is the lack of an underlying theoretical conceptualization. I had that uncomfortable feeling one gets whenever reading explanations of regression equations which show results--statistically significant, no less--but which are difficult to explain because there is no *a priori* theory, or meaningful hypotheses. The author has some underlying hunches, as evidenced by including F-scale measures and self-confidence items in the equation, but these need to be articulated and used as a basis for some hypotheses, or at least, expectations for the data.

### 3. Relevance to Practice

Now let me turn to *attention to relevance to management practice and/or to public policy issues*--the third criterion I suggested to evaluate whether studies in this area hold the long-term promise of establishing an important sub-field of consumer research, or whether we are on the downside of the PLC. In my opinion, the Bjorklund and Bjorklund study exhibits the most thorough and interdisciplinary conceptualization of any of these papers, and the study shows a great deal of procedural rigor. However, it is also the paper that gives me the most trouble in trying to determine its relevance. There may be some relevance to toy companies which use observational techniques to assess the amount of time children play with prototype toy products, but this possibility is not discussed by the authors. They imply that toddler satisfaction with their toy environment has an effect on consumer socialization processes later, but there is no discussion of exactly how this might occur, or how their data bear on the process. Bjorklund and Bjorklund do cite studies which relate play satisfaction to developmental outcomes and cognitive performance, but they do not say what these developmental outcomes are, or how they might relate to consumer socialization processes. Finally, I have some difficulty with the assertion that time spent with a toy are valid measures of satisfaction: there is a tautology here, it seems to me. Admittedly, toddlers can't verbalize "satisfied" attitudes, and even among adults, conceptualization and measurement of consumer satisfaction is currently a major interest in consumer research. So perhaps we can't be too critical of the satisfaction measure, although one might suspect that the novelty of a cornucopia of toys boggled the little kids, and it might have been better to observe play time and behavior *after* children had been given a chance to familiarize themselves with the toys.

Jay Lindquist broaches--but does not quite fully address--a potentially important marketing management and public policy issue. In the current regulatory climate, the thought has occurred to many advertisers who target messages against child audiences to get "out of the kitchen" and move some of their media dollars to advertising vehicles other than television. Assuming that children's attitudes toward advertising in these vehicles is related to advertising effectiveness, managers should be interested in Lindquist's study. But just think how much more interested they would be if there were some information on media patterns among children in the age ranges surveyed, i.e., some information to help the advertiser evaluate the trade-off between reach and cost efficiency using television advertising, versus the possible credibility gains in using print media. In this study, for example, some available data from commercial sources would have been helpful to give readers a feeling for how many children read what kind of magazines, comics, etc., versus viewing and listening patterns. It also would have been helpful if the author had established

whether or not children ever heard of, read to some extent, and like, any or all of the publications (or other publications) they were shown as examples in class.

But let's get back to the managerial question of whether there are effective and efficient alternatives to television advertising to reach and influence kids. The overall finding in Lindquist's study is that children are most "negative"--in terms of overall attitude--toward television advertising. But a danger in using total attitude scores (TOTATT) is that some of the most interesting and useful information is obscured. If you look closely at Table 6, at the components of attitude, you find something interesting. Take sixth-graders, who are the best readers, have the most money to spend, and, perhaps, influence some parental purchases more than younger kids. It turns out that these children (and, in fact, kids at all grade levels) feel TV advertising more objectively describes advertised products than child magazine advertising; also, sixth graders feel TV advertising is more persuasive, and there's not much difference in their attitudes about believability and trustworthiness. If an advertising manager looked beyond the main reported result, he might decide that it isn't so hot in the kitchen after all.

There is a double-edged sword here in that too little attention has been devoted to measure development and testing, and it's useful to provide further testing of Rossiter's attitude scale. But the availability of measures sometimes blinds us to the most interesting and useful aspects of phenomena, while it steers us away from the tedious work involved in measure development. In this case, the most compelling data would seem to relate to how children process information in advertising in different media. Beyond attitudes, do children select, evaluate and use information differently, depending on the medium in which it appears? There is a good deal of literature in this area, and one hopes that, in future work, this literature will be explored and used as a conceptual basis for more explicit analysis of children's processing of advertising information in various media. This would help to form a basis for hypotheses, which would avoid the problem in this study that there is little basis for explanation of results. There are some hunches, but it is far more compelling to articulate hypotheses on the basis of an underlying conceptual framework than to offer explanations *post hoc*.

#### A Re-evaluation and Future Directions for Research

What can we learn from these studies about the shape and direction of knowledge about marketing and children? With the exception of Nelson's study, the papers here are all billed as preliminary or exploratory work. Consumer research has been criticized for having so much work in the "exploratory" stage, but anyone who has tried to do research with kids knows there are some unique pitfalls not encountered in more established modes of research with sophomores or with adults. It is advisable to go slowly, and the criticism is only valid if exploratory efforts are not followed by more definitive studies.

All of these studies *pay attention to previous research*. The Bjorklund's are particularly creative in reviewing and using work outside of marketing; Lindquist, Nelson and Jenkins each employ measures and frameworks used in previous research. There are problems in doing this, and I've tried to point out, but at least these authors have attempted to blunt another criticism often made of consumer research: that we do not pay enough attention to developing standard and reliable measures, and that we ignore useful frameworks in previous research.

There is also much to be said for the *specific focus* of the studies here. Jenkins and Nelson look *in detail* at children's influence on specific family decisions, and Lindquist focuses on a compelling management question. But the feeling I have is sort of like watching potters at work, shaping objects by holding them up against one of those big round wheels. These studies are just "one turn" away from the kinds of finished products that can really make substantive contributions to the field. In some cases, the "final turn" involves conceptual work, in others, greater attention to appropriate methodologies to the research questions.

For example, a recurrent problem concerns the reliability and validity of data from children--especially data based on verbal responses. The younger the kids, the worse the problem. The choices are to observe kids in some natural or artificial state, or to ask them questions. On the other hand, we can ask parents to recollect and reliably report things about kids. The papers here illustrate all of these approaches, and the authors candidly acknowledge the problems attending each of them. The choice of a research approach depends on careful delimitation of the problem for research. For example, the question of children's influence on parents' decisions suggests that we ask parents some questions. The more specific the questions, and the more discreet and objective the response categories, the better the data. Are we interested in the *process* of influence? What aspects of the process? Parental responses to kinds of direct and indirect requests by kids, or buying decisions parents make for kids regardless of their requests? Is the interest in the *range of goods and services* children influence? Or is the interest in the independent variables which predict different frequencies, types, or processes by which children influence parental purchases?

What about research problems which demand that we gather data from kids themselves? The issue is really not experiment versus survey, or verbal versus behavioral measures. The issue again seems to me to be the problem you're interested in. In some of our current research, we are interested in children's processing of information in television advertising differing in the type of appeal used, and the degree of repetition (Wackman, et al, 1977). So we have asked children questions, but before we did, we tried to plan the research so we had an *a priori* conceptual framework to use in coding the open-ended responses. We hope the results will be pertinent to policy decisions about types of appeals which foster or inhibit children's learning from commercials, and the further applied question of how educational materials could be designed to help children understand, evaluate and use advertising content better.

To sum up, the methodological choices are pretty straightforward once the objective of the research is set. And my belief is that the objective ought to be related to some identifiable decision alternatives. The studies here are pertinent to management decisions such as media scheduling and appeal formulation, depending on advertising effectiveness in different media reaching kids, or depending on what we can tell managers about how kids affect parental buying decisions in families. And these questions are relevant to public policy questions too, since they bear on the issues of the extent to which children bug parents to buy things, and the nature of parental mediation of children's influence--as Nelson and Jenkins show, its a complex process, indeed.

It's my feeling that attention to relevance distinguishes consumer research from more discipline-oriented research. Moreover, the most useful research proceeds from a basis of conceptual clarity. To the extent clarity and relevance guide the choice of research questions concerning marketing and children, I feel this sub-area of research can be a productive and methodo-

logically sound on-going tradition in consumer research,  
rather than a temporary rash of exploratory studies.

#### References

A.R. Andreasen, Remarks at Tenth Annual Paul D. Converse Awards Symposium, University of Illinois, Champaign-Urbana, Illinois, May, 1978.

E. Zigler and I. Child, "Socialization," in G. Lindzey and E. Aronson (Eds.), *The Handbook of Social Psychology*, Vol. 2 (Reading, Mass., Addison-Wesley, 1969).

W.D. Wells, "Communicating with Children," Journal of Advertising Research, 5 (June, 1965), 2-14.

D.B. Wackman, E. Wartella and S. Ward, "Children's Information Processing of Television Advertising," National Science Foundation Grant, 1977.

CONCEPTUAL AND METHODOLOGICAL ISSUES  
IN CONSUMER SATISFACTION RESEARCH<sup>1</sup>

Stephen A. LaTour, University of Florida  
Nancy C. Peat (student), University of Florida

Abstract

Consumer satisfaction has recently become a particularly salient concern for both business and government. This increased interest, however, has generated a relatively meager amount of research designed to assess the determinants of consumer satisfaction and dissatisfaction. This paper critically analyzes the existing survey and experimental research in this area, focusing on conceptual and methodological issues. An alternative model of consumer satisfaction based upon Thibaut and Kelley's (1959) comparison level theory is proposed.

Introduction

Consumer satisfaction has long been a central concern of modern marketing practitioners (Kotler, 1976) and more recently a major concern of various governmental agencies engaged in consumer protection activities. The consumerist movement has made consumer satisfaction an even more salient concern for both business and government by calling attention to consumer dissatisfaction with products, services, and marketing practices.

Unfortunately, this increased interest in and concern about consumer satisfaction/dissatisfaction has generated a relatively meager amount of research designed to unravel the determinants of consumer satisfaction/dissatisfaction. The research can be classified into two major streams: 1) survey research designed to assess a) incidences of satisfaction/dissatisfaction for various products and services (e.g. Handy and Pfaff, 1975), and b) these incidences together with an assessment of the relationship between satisfaction and various socioeconomic and demographic characteristics (e.g. Handy, 1977; Ash, 1978) and 2) mostly experimental research designed to assess the impact of expectations about product performance on post-consumption satisfaction or perceived product performance (e.g. Cardozo, 1965; Olshavsky and Miller, 1972; Anderson, 1973). The purpose of this paper is to critically evaluate the conceptual and methodological problems of this research and to propose an alternative conceptualization of the determinants of consumer satisfaction based upon Thibaut and Kelley's (1959) comparison level theory.

Critical Review of Consumer Satisfaction Research

Survey Research

Survey research has revealed that degree of consumer satisfaction/dissatisfaction varies widely among individuals as well as over product and service categories. Some survey researchers have attempted to account for differences in satisfaction by correlating consumer satisfaction/dissatisfaction with an extensive variety of socioeconomic and demographic variables. The usefulness of this "shotgun approach" (c.f. Kassirjian, 1971) is questionable, not only because it capitalizes on chance, but also because it lacks a clearly stated conceptual rationale. In addition, the statistically sig-

nificant relationships that have been obtained usually account for relatively small percentages of variance in consumer satisfaction (c.f. Grønhaug, 1977; Day and Bodur, 1977; Ash, 1978). Thus while survey research has provided some useful descriptive data, it has not contributed very much to our understanding of the determinants of consumer satisfaction/dissatisfaction.

Research on Consumer Expectations

Some consumer researchers have theorized that satisfaction is a function of the discrepancy between a consumer's expectations about the performance of a product and obtained product performance. A number of functional relationships between expectations and obtained performance have been proposed to explain satisfaction, but the research has concentrated on three basic approaches.<sup>2</sup> First is Howard and Sheth's (1969), Engel, Kollat and Blackwell's (1973), and Cardozo's (1965) contrast hypothesis. "Contrast theory," traceable to Helson's (1964) adaptation-level theory, suggests that consumers will compare actual performance of a product to their expectations about performance. If obtained performance is less than expected, consumers will be dissatisfied. On the other hand, if their expectations are met or exceeded, they will be satisfied.

The second approach is derived from assimilation-contrast theory (Sherif and Hovland, 1961). According to this theory, expectations serve as an anchor for the judgment of product performance, and assimilation or contrast effects will occur as a function of the degree of discrepancy between obtained and expected performance (c.f. Anderson, 1973, Olson and Dover, 1976). If the discrepancy is not too large, assimilation should occur. In this event expectations higher than obtained outcomes should result in higher performance judgments (and therefore greater satisfaction) than if expectations match obtained performance. Similarly, expectations lower than obtained outcomes should result in perceptions of poorer performance (and less satisfaction) than if expectations match obtained outcomes. If, however, the discrepancy is very large, a contrast effect should occur; expectations higher (lower) than obtained outcomes should result in lower (higher) performance judgments (and consequently, less (more) satisfaction) than if expectations match obtained outcomes. This theory may be viewed as a refinement of "contrast theory," for contrast effects are hypothesized to occur only when the discrepancy between expected and obtained performance is extreme.

The third approach, based upon various psychological consistency theories (c.f. Olshavsky and Miller, 1972; Anderson, 1973), hypothesizes that consumers will experience psychological tension when an inconsistency exists between an expected and obtained level of performance. The unpleasantness generated by this tension will cause consumers to (mis)perceive product performance as consistent with their expectations. Thus as expectations increase, perceived product performance should increase

<sup>1</sup>The authors wish to express their appreciation to Olli Ahtola and Pauline Houlden for their helpful comments and suggestions.

<sup>2</sup>Other functional relationships such as those based upon the theory of generalized negativity (Carlsmith and Aronson, 1963) have been suggested, but these generally have not been replicated and will not be discussed.

and greater satisfaction should result.

A few investigations have been conducted to test the validity of these models of consumer satisfaction. Unfortunately, rather than measuring the impact of expectations on satisfaction, they have usually measured the impact of expectations on perceived product performance or quality (Cardozo, 1965; Anderson, 1973; other studies such as those by Olshavsky and Miller (1972) and Olson and Dover (1976) have been misclassified by various researchers as studies of the relationship between expectations and satisfaction when in fact they only claim to study the relationship between expectations and perceived performance). The investigators have simply assumed that assessment of perceived product performance on quality is equivalent to assessing satisfaction; however, perceived product performance or quality is not necessarily the same as satisfaction. While perceived product performance is likely to be highly correlated with amount of satisfaction, there are other variables that may interfere with this relationship. For example, a product or service provider may perform well on a number of dimensions but the product or service may cost too much or be too distant from the consumer. In such instances the consumer could be dissatisfied. In addition, unless performance measures contain an affective component (e.g. on this dimension the product performs poorly, neither poorly nor well, well), they will not indicate whether the consumer will be satisfied or dissatisfied with the obtained level of performance. For example, Anderson (1973) measured perceived quality by having subjects estimate the price of a pen they evaluated. There is no method of determining from this data what level of quality would be required for the consumer to describe himself or herself as satisfied rather than dissatisfied. In sum, care must be taken in making inferences about satisfaction on the basis of quality or performance judgments. Nonetheless, a careful examination of the research concerning the effect of expectations on perceived performance is in order because the amount of variance in satisfaction accounted for by product quality or performance is likely to be fairly high. Unfortunately, examination of this research reveals serious methodological problems which cause further difficulties in making inferences about the effect of expectations on consumer satisfaction.

For example, Cardozo (1965) found contrast effects when subjects were asked to evaluate desirability and quality of ballpoint pens. But closer examination of his results indicates that anchors for the dependent measures varied as a function of whether expectations were high or low (c.f. Olshavsky and Miller, 1972). Thus the validity of Cardozo's contrast effect may be questioned and the effect of expectations on satisfaction is not clarified by this study.

Olshavsky and Miller (1972) found that higher expectations produced higher performance judgments and claimed support for a consistency theory interpretation of the results. Analysis of their methodology suggests, however, that this interpretation is untenable. Olshavsky and Miller simply asked subjects to evaluate a product presented to them. In this circumstance, degree of involvement with the task would have been extremely low since subjects made no choice among products and their outcomes were not affected by the quality of the product as they would have been in a true purchase situation. Thus it is unlikely that the tension necessary to create consistency effects would have been established. The results might be interpreted as an assimilation effect (which does not postulate the tension required by consistency theories); however, further examination of their methodology indicates that even this interpretation is questionable. In their study different expectations were created about a tape recorder's performance by varying the ratings contained in an authoritative test report. After listening to the tape recorder described

in the test report, subjects were asked to judge the performance of the recorder. Since subjects may have interpreted an instruction to rate the recorder as an instruction to rate the brand of recorder in general rather than the tested sample, they may have relied more heavily on beliefs created by the test reports than they would have otherwise. It may have been particularly important to rely upon the test report in order to make an accurate judgment since subjects listened to only one tape and may have been concerned about avoiding judgmental errors resulting from temporary factors such as the quality of the tape, the competency of the person making the recording, or a defect in this particular sample of the brand. In addition, the experiment may have created demand characteristics (Orne, 1962). There may have been an implicit demand to use the test report information in arriving at a judgment since subjects probably perceived that the experimenter intended them to use the supplied information (i.e. the purpose of the experiment was probably transparent to the subjects). The probable existence of evaluation apprehension (Rosenberg, 1969) may have enhanced the impact of these demand characteristics. Subjects would have wanted the experimenter to evaluate their judgments favorably and therefore may have felt further compelled to use the information supplied by the experimenter. Thus possible judging of the brand rather than the tested sample, demand characteristics, and evaluation apprehension may have all tended to exaggerate the influence of the expectation-producing information on subjects' judgments of product performance. Hence the observed assimilation effect may have been artificial.

Interpretation of the findings of Anderson's (1973) study is also difficult. Subjects were first given information which created differential expectations about the quality of a pen and were then asked to evaluate the pen. Anderson interpreted his results as demonstrating an assimilation effect because expectations above the objective quality level produced higher quality judgments than expectations matching obtained quality, while expectations below the objective level produced lower quality judgments. Anderson also perceived what he labeled a contrast effect -- when expectations were extremely high, judgments of product quality were somewhat lower than when expectations were moderately high. This, however, is not a true contrast effect since such an effect would require perceived quality to be less than objective quality.<sup>3</sup> As in Olshavsky and Miller's study, subjects may have interpreted the instructions to evaluate "this pen" as a request to evaluate "this brand of pen" rather than "this particular sample." In addition, they may have felt an implicit demand to use the expectation-producing information provided by the experimenter and the effect of this demand may have been magnified by the existence of evaluation apprehension. Again, the assimilation effect may well have been artificial.

Olson and Dover's (1976) investigation reveals a possible assimilation effect for judgments of coffee bitterness but the results are not statistically significant. The authors argue that one of their four measures of bitterness showed a significant difference between experimental and control groups but the probability level was only .10. This is marginal at best, given the lack of a multivariate test or some other method of guarding against Type I errors when multiple univariate tests are employed for conceptually similar dependent variables. In addition, while demand characteristics were eliminated in this study, subjects were still asked to rate the brand of coffee, not the particular sample they tested. They may have believed the sample was bitter (and would have been dissatisfied with it) but because they were asked

<sup>3</sup>The relative ineffectiveness of the extremely high expectation was probably due to subjects attaching less weight to information lacking in credibility.



to evaluate the brand, they may have relied less on the sample they tasted -- which could have been influenced by temporary factors -- than on the expectation-producing information supplied by the experimenters. Thus if one is interested in the effect of expectation-producing information on the consumer's response to a given purchase experience, even this small trend toward an assimilation effect may be artifactual.

Finally, Swan and Combs (1976) measured both expectations about and satisfaction with clothing purchases. They found that consumers who were dissatisfied with clothing purchases were more likely to have had their expectations negatively disconfirmed than were consumers who were satisfied with purchases. This seems to support the contrast hypothesis. Nonetheless, before accepting this interpretation of the effect of expectations on satisfaction, it should be noted that this study was based on recall of particularly satisfying and dissatisfying purchases. Thus there may have been some retrospective distortion of expectations. In addition, subjects were not asked to report middle-range experiences for which assimilation effects might have been found. If this had been done, assimilation-contrast theory might have been supported instead.

Given these problems, what can be said about the effect of expectations on satisfaction? We would speculate that expectations will have some influence on the consumer's post-consumption beliefs about the performance of a product on various attribute dimensions but only when the consumer is uncertain about the level of the attribute possessed by the product (i.e. the dimension is ambiguous). In such cases an assimilation effect should occur. (The attributes in the studies discussed above were not very ambiguous, and the fairly strong assimilation effects for judged performance or quality were probably the result of methodological problems -- particularly demand characteristics. In Olson and Dover's (1976) study where the possibility of demand characteristics was reduced, the size of the assimilation effect was very small.) A contrast effect for post-consumption beliefs about the performance of a product on an attribute has never been demonstrated. Indeed, no matter how outlandish a claim by a manufacturer, we would not expect the consumer to think the product had less of an attribute than that judged to exist in the absence of a manufacturer's claims. When manufacturer's claims are extreme, the consumer would attach less weight to them and they would not influence the consumer's belief about an attribute as much as they would had the claims been moderately high. Thus for ambiguous dimensions, expectations above the objective performance level may increase judgments about the quality or performance of the product and therefore increase satisfaction. But, even though expectations above objective performance levels may increase judgments about the level of an attribute believed to be possessed by a product, there will still be some discrepancy between the level believed to be obtained and the expectation level (unless the attribute dimension is totally ambiguous). This discrepancy would not have occurred had expectations been veridical with the objective attribute level. According to contrast theory, this discrepancy which remains between the subjective and expected attribute levels will contribute to dissatisfaction. Thus the gain in satisfaction obtained by increasing perceived quality through creation of high expectations might be eliminated by the dissatisfaction associated with this discrepancy. Obviously, the more ambiguous the attribute, the more the consumer will be forced to rely on the expectation-producing information rather than personal judgment in forming a belief about the attribute level. The more ambiguous the attribute, the smaller the discrepancy between the subjectively obtained level and the expected level, and thus the more effective high expectations would be in producing greater satisfaction. Clearly, if dimensions are unambiguous, expecta-

tions above objective levels of performance would not produce any increase in perceived quality and the result of such expectations would be consumer dissatisfaction. Overall, we would hypothesize that the creation of expectations above objective performance or quality levels would typically result in some degree of consumer dissatisfaction.

#### Further Problems and Issues

##### Expectations -- the Primary Determinant of Consumer Satisfaction?

In addition to methodological and inferential problems with the research on consumer expectations, one may well question the theory that consumer satisfaction or dissatisfaction with products and services results primarily from the confirmation or disconfirmation of expectations created by manufacturers' or service providers' claims.<sup>4</sup> Let us begin by clearly specifying what is meant by expectations. As Olson and Dover (1976) have noted, researchers have been less than precise in defining this construct; however, as our review indicates, their manipulation of expectations reveals that researchers construe them to be pre-decisional or pre-consumption beliefs about the overall performance of the product created by manufacturers' claims or test-report information. Arguing from the perspective of Ahtola's (1975) multiattribute attitude model, Olson and Dover conclude that expectations should be defined as the consumer's beliefs about the levels of attributes possessed by a product. This is an approach which we favor because of its high degree of specificity and diagnostic usefulness. Both of these approaches fail, however, to recognize the role of the consumer's affective response to obtained attributes in determining satisfaction. For example, a consumer may be forced to purchase a brand whose attributes he or she does not evaluate favorably (e.g. the consumer may have moved to a new market area or the manufacturer may have discontinued marketing the product whose attributes the consumer liked). Even though the consumer's predecisional expectations (beliefs) about the attributes of the brand may be confirmed by post-purchase experience with the product, the consumer will still be dissatisfied with the product because of his/her unfavorable evaluation of its attributes.

The disconfirmation of expectations model would fail to account for consumer satisfaction/dissatisfaction in other circumstances as well. A new brand may be produced with attributes the consumer evaluates more favorably than the attributes of other brands presently on the market. Even if the manufacturer created expectations that were too high, the consumer would not be dissatisfied, for despite the negative disconfirmation of expectations, the brand possesses more of the desired attributes than did previous brands. This is not to deny the role that disconfirmation of expectations may play in determining consumer satisfaction. Indeed, it would be reasonable to expect that the consumer in the first example (forced to purchase an unsatisfactory product) might have been more dissatisfied had the disliked attributes been unexpected, and the consumer in the second example (purchasing a product with favorable but less-than-expected attributes) would have been more satisfied had the manufacturer not created unrealistic expectations. We would speculate, however, that the consumer's evaluation of the product's attributes may account for more of the variability in post-purchase

<sup>4</sup> Expectations can also be the result of past experience with similar products or knowledge about other consumers' experiences with similar products. Later sections discuss the central role these experiences may play in determining satisfaction/dissatisfaction.

satisfaction than would the confirmation or disconfirmation of expectations about those attributes.

#### Discriminant Validity of the Satisfaction Construct

A further issue for consumer satisfaction researchers is the discriminant validity of the satisfaction construct. Satisfaction with a product is simply an evaluative response to that product. (Although some researchers may be more interested in satisfaction with the purchase circumstances, that is also an evaluative response.) Yet, a different construct, attitude, has been more frequently employed by consumer behavior researchers to refer to evaluative responses to products. Given that attitude and satisfaction are both evaluative responses to products, it is not clear whether there are any substantive differences between the two. In fact, it may be more parsimonious to consider satisfaction measures as post-consumption attitude measures. This does not mean that satisfaction is an inappropriate measure of the post-decision evaluative response or that it should be supplanted by other measures used by attitude researchers. Rather, this implies that we should be careful about attaching any special significance to satisfaction as a construct when it may simply be one measure of an attitude.

If satisfaction with a product is essentially the same as attitude toward a product, one might ask why satisfaction research has not been integrated with attitude theory and research. In addition, if the two concepts are highly similar, one might ask whether theory and research under the label of attitudes might contribute to an understanding of theory and research under the label of satisfaction, and vice versa.

Perhaps one of the major reasons for lack of integration of the two areas is that attitude researchers have focused largely on pre-decisional evaluations of products while satisfaction researchers have focused largely on post-consumption evaluative responses. In addition, satisfaction researchers have focused on the confirmation/disconfirmation of expectations (beliefs) about attributes while attitude researchers (particularly those interested in multiattribute models) have focused on expectations (beliefs) about attributes, evaluations of those attributes, and in some instances the importance of those attributes in determining the consumer's attitude (see Fishbein and Azjen (1975) for a discussion of the merit of importance weights). It is understandable that pre-decisional attitude research has not focused on the confirmation or disconfirmation of expectations, but it is not clear why satisfaction researchers have largely ignored the role of evaluations associated with obtained attributes. Perhaps it is time for satisfaction researchers to give more consideration to the role of these evaluations.

#### An Alternative Conceptualization of Consumer Satisfaction

Although many researchers in the area of consumer satisfaction are not fully acquainted with Thibaut and Kelley's (1959) comparison level theory, it may provide a useful basis for understanding post-consumption satisfaction/dissatisfaction with products and services. A modification of this theory, described below, may also provide insight into the determinants of consumer satisfaction with each of a product's component attributes, some combination of which presumably determines the consumer's overall satisfaction with the product.<sup>5</sup>

Thibaut and Kelley are social psychologists who concep-

<sup>5</sup>Miller (1977) appears to be the only satisfaction researcher who acknowledges the usefulness of their approach.

tualize the result of any interaction in terms of rewards and costs. In the consumer setting these rewards and costs might include (but would not necessarily be limited to) the attributes of the product or service (including its cost), the pleasures or perhaps difficulties associated with interpersonal interactions with salespeople or service providers, and the approving or disapproving responses of other persons to the purchased product. The average of obtained rewards and costs constitutes the outcome of an interaction. Clearly, some researchers might be interested in rewards and costs associated with the product itself while others might be more interested in the rewards and costs attendant to the circumstances surrounding the purchase experience.

According to Thibaut and Kelley, the consumer's satisfaction with the outcome of a purchase would be determined by the discrepancy between the outcome and a standard of comparison known as the Comparison Level (CL). Outcomes above the CL (positive discrepancies) will be satisfying while those below the CL (negative discrepancies) will be dissatisfying.<sup>6</sup> The comparison level itself is thought to be determined by the average of salient outcomes for the same or similar interactions that one has experienced or is aware of.<sup>7</sup> Outcomes that one is aware of but has not personally experienced would include known outcomes obtained by similar others (see Festinger's (1954) discussion of social comparison theory) and to a lesser extent outcomes one expects to obtain from the interaction (where these expectations are based on the unique characteristics of the present interaction rather than previous experience or the known outcomes of other people). Thibaut and Kelley's discussion suggests that these outcomes are weighted in the following order to determine the CL: first and most importantly, outcomes a person has directly experienced; second, outcomes of similar others; and third, special expectations created by the present interaction. Drawing a parallel to the situation of the consumer, this means that the comparison level for the outcome associated with any given product or service will be a function of experiences with similar products or services in the past, to a lesser extent the experiences of similar consumers, and to a still lesser extent outcomes promised by the manufacturer, retailer, and/or service provider.

Thibaut and Kelley's theory is particularly appealing because it explicitly recognizes that satisfaction is not an absolute phenomenon but rather a relative one. (Although the disconfirmation of expectations model also assumes that satisfaction results from a relative comparison of obtained attributes to expected attributes, as outlined above, the model is too restricted to account fully for consumer satisfaction.) Thus if consumers purchase a product with attributes they evaluate more positively than those obtained from previously purchased products, they will be satisfied (assuming outcomes of similar consumers and expectations specific to the purchase situation are held constant). For example, consumers who are used to tires that last 5,000 miles will be satisfied with tires that last 10,000 miles while consumers who are used to tires lasting 30,000 miles will be dissatisfied with those same tires.

<sup>6</sup>It should be noted that Thibaut and Kelley also consider the CL to be the deserved level of outcomes. It seems to us, however, that the deserved level would at least be minimally satisfying and would therefore be above the affectively neutral point on the outcome scale.

<sup>7</sup>Thus the concept of comparison level is similar to Helson's (1964) concept of adaptation level except that the comparison level is a standard for determining evaluative responses while the adaptation level is usually considered to be a standard for non-evaluative judgments (e.g. the brightness of a light).

While this approach to understanding overall satisfaction with a product or service appears to be useful, a word of caution is in order. Thibaut and Kelley argue that satisfaction is determined by averaging the rewards and costs associated with an interaction and comparing the result with the comparison level. However, Thibaut and Kelley define rewards as "the pleasures, satisfactions, and gratifications the person enjoys" (p. 12, emphasis ours). Their definition of costs employs a somewhat different terminology but suggests that costs are unpleasant or dissatisfying aspects of the interaction. If rewards are merely satisfactions and costs merely dissatisfactions, then the satisfaction associated with the obtained outcome is the average of the component rewards and costs and recourse to a comparison level is unnecessary to determine overall satisfaction with that outcome.

#### A Modified Comparison Level Theory of Satisfaction

While the comparison process described above may therefore be meaningless because the obtained outcome already contains the necessary affective information, extension of the concept of comparison to the attribute level may prove to be meaningful. We would propose that for each attribute there is a comparison level. This comparison level is a function of past experience with various levels of the attribute (for similar products used for similar purposes), levels of the attribute the consumer is aware that similar consumers receive, and expectations created by the manufacturer, salesperson, and/or service provider. These attribute levels are not measured in terms of rewards and costs but rather in terms of natural scale values (e.g. size, number of miles per gallon, number of compliments received for a new dress). For example, a consumer's satisfaction with the miles per gallon achieved by his/her new car would probably be related to the discrepancy between obtained MPG and a weighted average of his/her previous cars' average MPG, the average MPG his/her friends obtain from their cars, and the manufacturer's promises about the MPG he/she should obtain from the car. It can be speculated that the consumer's overall satisfaction with a car would be an additive function<sup>8</sup> of the discrepancies from the comparison level for each salient attribute (e.g. MPG, size, trunk space, braking distance, riding comfort) of the car. In determining overall satisfaction, each discrepancy would be weighted by the importance a consumer attaches to that attribute dimension. In addition, the discrepancies should be thought of as percentages of the comparison level to eliminate scaling differences among the various attribute dimensions. (This can be accomplished by dividing the discrepancy by the comparison level.)

For the attributes where more of the attribute is always desirable (a positive, infinite ideal point, e.g. MPG, tire mileage), positive discrepancies should contribute to satisfaction while negative discrepancies should contribute to dissatisfaction. For attributes where less of that attribute is always more desirable, the effect of discrepancies will be reversed. If an attribute dimension involves a finite ideal point<sup>9</sup> (e.g., the sweetness of a soft drink), discrepancies from two comparison levels must be considered in determining the satisfaction associated with that attribute: one below

<sup>8</sup>It is conceivable that this could be an averaging function; however, it is not the purpose of this paper to further explore the adding-averaging controversy.

<sup>9</sup>Finite ideal points are likely to be complexly determined; there may be some physiological determinants, a particular attribute level may have been associated with happy or sad events (classical conditioning), and/or an attribute level may have been the one most frequently experienced (familiarity).

and one above the ideal point. The comparison level below the ideal point can be defined as a weighted average of the experienced, known, and promised levels of attributes below the ideal point while the comparison level above the ideal point will be a weighted average of experienced, known, and promised levels of attributes above the ideal point. Obtained attribute levels below the ideal point would be compared to the lower CL while obtained attribute levels above the ideal point would be compared to the higher CL. As long as an outcome is between the two comparison levels, a consumer will be satisfied with the attribute level. (Similar logic could be applied to a finite negative ideal point except that obtained attribute levels between the two comparison levels would be dissatisfying while all other obtained attribute levels would be satisfying.) For example, with respect to a positive, finite ideal point, a consumer might ideally prefer a soft drink to be "moderately sweet." If the consumer has always obtained this outcome level, then the ideal point would define a single comparison level and any level of sweetness above or below the ideal point would be at least somewhat dissatisfying. It is unusual, however, for a consumer to have experienced only the ideal level of outcomes. Thus some levels below and above the ideal point will probably have become slightly or moderately satisfying (though certainly not maximally satisfying as the ideal point would be). The average sweetness experienced, known, or promised below the ideal would determine the lower comparison level. Sweetness between this level and the ideal point would be described by the consumer as satisfying. The average sweetness experienced, known, or promised above the ideal would determine the upper comparison level. Sweetness between the ideal point and the upper comparison level would also be described as satisfying by the consumer. All other levels of sweetness would produce varying degrees of dissatisfaction, depending upon their discrepancy from the appropriate comparison level. Overall satisfaction with a soft drink should be a sum of the discrepancies of all relevant attributes from their appropriate CL's with each discrepancy weighted by the importance of the attribute with which it is associated.

The proposed model can be formally expressed as:

$$S_j = \sum_{i=1}^n I_i (A_i - CL_i) / CL_i \quad (1)$$

$$= \sum_{i=1}^n I_i \left[ \left( A_i - \left\{ \frac{w_1 P_i + w_2 R_i + w_3 E_i}{w_1 + w_2 + w_3} \right\} \right) / \left( \frac{w_1 P_i + w_2 R_i + w_3 E_i}{w_1 + w_2 + w_3} \right) \right] \quad (2)$$

where:

- $S_j$  = post-consumption satisfaction with brand j
- $I_i$  = importance the consumer attaches to attribute i
- $A_i$  = subjectively experienced attribute level
- $CL_i$  = comparison level for attribute i
- $P_i$  = average of personally experienced attribute levels
- $R_i$  = average of attribute levels experienced by referent persons (similar others)
- $E_i$  = expected value of the attribute level based on the unique characteristics of the present purchase situation (i.e. expectations orthogonal to  $P_i$  and  $R_i$  such as expectations created by manufacturer's claims)

$n$  = number of salient attributes

$w_1, w_2, w_3$  = empirically determined weights

It should be noted that this formula assumes a positive, infinite ideal point for each attribute. For attributes having a negative, infinite ideal point, the sign of the

discrepancy is the same as in the formula for outcomes below the ideal point; but for outcomes above the ideal point, the sign must be reversed. For negative, finite ideal points, the signs of the obtained discrepancies would be reversed from what they would be for a positive, finite ideal point. Of course, the lower or upper CL should be employed depending upon whether the obtained level of an attribute is below or above the ideal point. While the formula assumes constant weighting of  $P_i$ ,  $R_i$ , and  $E_i$  across the attributes, this assumption may need revision as the model is further explored. Finally, note that because the model assumes  $A_i$  is the subjectively experienced attribute level, it can account for situations in which the consumer is uncertain about the obtained level (i.e. where the dimension includes some ambiguity). In such cases the subjectively experienced attribute level is defined as the expected value of the attribute level. That is,  $A_i$  would be an additive function of the subjective probabilities associated with each attribute level multiplied by the respective values of those levels.<sup>10</sup>

#### Implications of the Proposed Satisfaction Model for Attitude Theory

As indicated previously, attitude theory (in particular multiattribute theory) has implications for consumer satisfaction research because it reveals the necessity of considering the consumer's evaluative (satisfaction/dissatisfaction) response to obtained attributes. Discussion is now in order concerning the implications that the proposed model of satisfaction may have for attitude theory. In particular, the extension of Thibaut and Kelley's concepts to the level of the component attributes of a product provides unique insight into the determinants of the evaluative responses to those attributes. Attitude theorists (particularly those employing Fishbein's (1967) model or variants of his model) have typically assumed that the evaluation of an attribute level is the result of classical and/or operant conditioning. It has been demonstrated that this can happen (c.f. Fishbein and Ajzen, 1975), but whether this is the predominant method by which affect is attached to various attribute levels may be questioned. We would argue that attitudes toward various attribute levels are rarely classically conditioned. It is unusual for an attribute level to be associated consistently with some particular good or bad experience (although ideal points may in part be established by such experiences). In fact, if classical conditioning of responses to attribute levels were to occur very often, affect curves would probably be very jagged -- but this is a rare occurrence. Operant conditioning would also fail as a complete explanation of affective responses to attribute levels. For example, we could say that the less time a vacuum cleaner took to clean a carpet, the more a consumer would be reinforced by using it. But saying this begs the question concerning the point at which the evaluation of amount of time spent vacuuming changes from unfavorable to favorable (i.e. from dissatisfaction to satisfaction). We would argue instead that whether the amount of time is evaluated favorably will be a function of the amount of time one has taken in the past, the amount of time one is aware that similar others' vacuum cleaners take, and expectations created by the vacuum cleaner's manufacturer. Thus the process of judgment relative to a comparison level would seem to better account for the formation of attitudes toward attribute levels.

<sup>10</sup> A similar model could be developed to determine consumer satisfaction with the purchase circumstances by specifying the salient attributes of the purchase situation. In many circumstances it may be useful to incorporate both product attribute and purchase circumstance attributes into the model.

#### Cautions Concerning the Proposed Model

Four final caveats are in order concerning the proposed model of satisfaction. First, consumers will not demonstrate infinite adaptive capabilities. Even if a consumer has experienced only perfumes with the pungent odor of rotten eggs, it is unlikely that such a consumer could ever be satisfied with such a perfume although he/she would probably be less dissatisfied with such a pungent smell than most consumers in our society. In sum there are probably physiological absolutes which have an impact upon affective responses to attributes. Nonetheless most product attribute levels are probably well within human adaptive abilities.

Second, it should be noted that the proposed model assumes the attributes to be measureable on a scale with metric properties. This requirement may appear to limit the usefulness of the model when an attribute dimension is of a qualitative nature. For example, with respect to applesauce, a consumer may be concerned about whether the container is metal or glass. These may appear to be qualitatively different attributes, but this simply means that metal and glass packages differ on more than one underlying dimension. The consumer has simply used the glass-metal distinction to summarize his/her concerns about these underlying dimensions. For example, in this case the underlying dimensions might be breakage resistance and amount of alteration in taste caused by the package. These dimensions are quantifiable and employable in the proposed model.

Third, in order for attribute dimensions to be effectively employed in the model, it must be assumed that there are no interactive effects of the attributes (an assumption made by virtually all multiattribute models). If an interaction occurs, the ability of the model to predict satisfaction will be diminished. Nonetheless, while there may be situations where this will occur, we would argue that the proposed model can account for a large proportion of the variability in consumer satisfaction. In addition, should the need arise, interaction terms can be included in the model.

Finally, we are not suggesting that this model be employed as a primary measure of consumer satisfaction for various products and services. Simple, direct measures of satisfaction are available for this purpose. What we have proposed is a conceptual model of the major determinants of consumer satisfaction which, if verified by empirical research, should provide further insight into the causes of consumer satisfaction/dissatisfaction.

#### Conclusions

Our analysis of consumer satisfaction research suggests that close attention to conceptual and methodological issues will be necessary if further progress is to be made toward understanding the determinants of consumer satisfaction. Hopefully, the proposed psychological model of satisfaction will contribute important insight into these determinants.

#### References

- Olli T. Ahtola, "The Vector Model of Preferences: An Alternative to the Fishbein Model," Journal of Marketing Research, 12 (February, 1975), 52-9.
- Rolph E. Anderson, "Consumer Dissatisfaction: The Effect of Disconfirmed Expectancy on Perceived Product Performance," Journal of Marketing Research, 10 (February, 1973), 38-44.

- Stephen B. Ash, "A Comprehensive Study of Consumer Satisfaction with Durable Products," in H. Keith Hunt, ed., Advances in Consumer Research, Volume V, (Ann Arbor: Association for Consumer Research, 1978), 254-62.
- Richard N. Cardozo, "An Experimental Study of Consumer Effort, Expectation and Satisfaction," Journal of Marketing Research, 2 (August, 1965), 244-9.
- J. Merrill Carlsmith and Elliot Aronson, "Some Hedonic Consequences of the Confirmation and Disconfirmation of Expectancies," Journal of Abnormal and Social Psychology, 66 (February, 1963), 151-6.
- Ralph L. Day and Muzaffer Bodur, "A Comparative 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.
- James F. Engel, David T. Kollat and Roger D. Blackwell, Consumer Behavior (second ed.), (New York: Holt, Rinehart and Winston, 1973).
- Leon Festinger, A Theory of Cognitive Dissonance, (Evanston, Illinois: Row, Peterson, 1957).
- Martin Fishbein, "A Behavior Theory Approach to the Relations Between Beliefs About an Object and the Attitude Toward the Object," in Martin Fishbein, ed.) Readings in Attitude Theory and Measurement, (New York: John Wiley & Sons, 1967).
- Martin Fishbein and Icek Ajzen, Belief, Attitude, Intention and Behavior, (Reading, Massachusetts: Addison-Wesley Publishing Company, 1975).
- Kjell Grønhaug, "Exploring Consumer Complaining Behavior: A Model and Some Empirical Results," in W. D. Perreault, ed., Advances in Consumer Research, Volume IV, (Atlanta: Association for Consumer Research, 1977), 159-165.
- Charles R. Handy, "Monitoring Consumer Satisfaction with Food Products," in H. Keith Hunt, ed., Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction, (Cambridge, Massachusetts: Marketing Science Institute, 1977).
- Charles R. Handy and Martin Pfaff, "Consumer Satisfaction with Food Products and Marketing Services," Agricultural Economic Report No. 281, Economic Research Service, U.S. Dept. of Agriculture, 1975.
- Harry Helson, Adaptation-Level Theory, (New York: Harper and Row, 1964).
- John A. Howard and Jagdish N. Sheth, The Theory of Buyer Behavior, (New York: Wiley, 1969).
- Harold H. Kassarian, "Personality and Consumer Behavior: A Review," Journal of Marketing Research, 8 (November, 1971), 409-18.
- Philip Kotler, Marketing Management: Analysis, Planning, and Control, (Englewood Cliffs, N.J.: Prentice-Hall, 1976).
- John A. Miller, "Studying Satisfaction, Modifying Models, Eliciting Expectations, Posing Problems, and Making Meaningful Measurements," in H. Keith Hunt, ed., Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction, (Cambridge, Massachusetts: Marketing Science Institute, 1977).
- Richard N. Olshavsky and John A. Miller, "Consumer Expectations, Product Performance, and Perceived Product Quality," Journal of Marketing Research, 9 (February, 1972), 19-21.
- 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, Volume III, (Chicago: Association for Consumer Research, 1976), 168-75.
- Martin Orne, "On the Social Psychology of the Psychological Experiment: With Particular Reference to Demand Characteristics and Their Implications," American Psychologist, 17 (1962), 776-83.
- M. J. Rosenberg, "The Conditions and Consequences of Evaluation Apprehension," in R. Rosenthal and R. L. Rosnow, eds., Artifacts in Behavioral Research, (New York: Academic Press, 1969).
- Muzaffer Sherif and Carl I. Hovland, Social Judgments: Assimilation and Contrast Effects in Communication and Attitude Change, (New Haven, Connecticut: Yale University Press, 1961).
- John O. Summers and Donald H. Granbois, "Predictive and Normative Expectations in Consumer Satisfaction and Complaining Behavior," in W. D. Perreault, ed., Advances in Consumer Research, Volume IV, (Atlanta: Association for Consumer Research, 1977), 155-8.
- John E. Swan and Linda Jones Combs, "Product Performance and Consumer Dissatisfaction: A New Concept," Journal of Marketing Research, 40 (April, 1976), 25-33.
- J. W. Thibaut and H. H. Kelley, The Social Psychology of Groups, (New York: Wiley, 1959).

## CONSUMER RESPONSE TO DISSATISFACTION WITH DURABLE PRODUCTS

Ralph L. Day, Indiana University  
Stephen B. Ash, University of Western Ontario

### ABSTRACT

Data were obtained from a sample of 119 households on instances of dissatisfaction, reasons for being dissatisfied, and the nature and extent of any subsequent complaining behavior for each of 63 categories of consumer durables. Substantial differences across categories were found in the fraction of users reporting dissatisfaction, reasons given for dissatisfaction, and subsequent action or inaction.

### INTRODUCTION

The literature of consumer satisfaction, dissatisfaction and complaining behavior is growing rapidly but still offers spotty coverage of the full range of products and services. Because unique aspects of particular products and services contribute to the wide variations in dissatisfaction which are observed over product categories, the need for comprehensive coverage is obvious. This paper presents some previously unreported results from a study which tested methods for collecting data in a single study over the full range of products and services. In an effort to develop a workable set of classes which would encompass all specific consumer products and services, about 200 categories were established. These were organized into three major sections; durables, non-durables, and consumer services. Some results have already been reported: analysis of satisfaction scale results for services (Day and Bodur, 1977); complaining behavior for services (Day and Bodur, 1978); and analysis of satisfaction scales for durables (Ash, 1978). This paper presents some analysis of the complaining behavior data from the durables section.

Most previous empirical studies have focused on the identification of correlates of consumer satisfaction, dissatisfaction, and complaining behavior. Early efforts were directed toward demographic correlates (Mason and Himes, 1973; Gronhaug, 1977; Warland et al., 1975; Liefeld et al., 1975; Thomas and Shuptrine, 1975). In recent papers, writers have directed their search for correlates of complaining behavior toward psychographic variables (Westbrook, 1977; Wall et al., 1977; Zaichkowsky and Liefeld, 1977). However a few studies have attempted to conceptualize and model the consumer complaint process (Day and Landon, 1977; Landon, 1977; Gronhaug, 1977). The framework for dealing with data on complaining behavior developed by Day and Landon (1977) and applied by Day and Bodur (1978) will guide the presentation of results below.

### OBJECTIVES

The original goal of the study was to design and test a research method to provide the kinds of information needed by public policy decision makers in monitoring and evaluating consumer satisfaction and dissatisfaction over the total range of consumer products and services (Day and Landon, 1976). However, as the study has evolved, the focus has shifted more toward providing descriptive data and insights which will be useful in conceptualizing the consumer evaluation process including any post-evaluation behaviors. This paper has the following specific objectives:

- (1) To compare levels of consumer dissatisfaction over various classes of durable products;
- (2) To determine which durable products appear to have caused the greatest amounts of dissatisfaction among users;
- (3) To identify recurring reasons for dissatisfaction with durable products;
- (4) To describe how consumers who report dissatisfaction attempt to resolve their complaints through alternative courses of private and public action;
- (5) To compare complaining behavior patterns over various classes of durable products.

Hopefully, study of these results will provide information and insights which will be useful in designing future research on complaining behavior.

### DATA COLLECTION

The data analyzed in this paper were obtained with self administered questionnaires using the drop off-pick up method to a probability sample of 600 dwelling units in Bloomington, Indiana during the fall of 1976 (Day and Bodur, 1977). The questionnaires obtain data on consumer satisfaction/dissatisfaction and complaint behavior on an "aided recall" basis. The product categories are divided into three major classes of goods and services and are utilized in four questionnaires: (1) non-durables; (2) durables; (3) services and intangible products; and (4) a combined version containing the entire 200 classes. Each of the three separate questionnaires further sub-divide the product or service categories with a full set of complaining behavior questions at the end of each subsection. The results reported here were obtained from data covering the 63 categories of durable products which were sub-divided into four classes: (1) housing and home furnishing durables; (2) home appliances and personal care items; (3) entertainment, recreation and education durables; and (4) cars and other transportation durables. A two-stage area sampling plan was employed to gather the data. The questionnaires were delivered by MBA and advanced undergraduate students enrolled in marketing research courses. The overall response rate was approximately 80%. A total of 119 usable separate durable products questionnaires were obtained.

The questionnaire first required respondents to indicate whether or not they had purchased any items from the category during the three year recall period. Those who had were then asked to provide an indication of the relative importance of that category and of their degree of satisfaction or dissatisfaction with items contained in it. At the end of each section, respondents indicated if they had been highly dissatisfied by items contained in any of the categories in it. Those reporting extreme dissatisfaction then indicated the number of times they had experienced it. They were then asked to identify the category which was most unsatisfactory and to report the specific name of the item. Those reporting dissatisfaction completed a set of questions identifying their

reasons for dissatisfaction and what steps were taken, if any, to resolve their dissatisfaction.

To facilitate discussion, the results are sub-divided into three classifications: (1) the number of instances of dissatisfaction; (2) reasons for dissatisfaction with durables; and (3) post-dissatisfaction responses.

#### Levels of Consumer Dissatisfaction

Tables 1 through 4 contain summaries of all durable product categories and show proportions of use and dissatisfaction for the various items. One can see wide variation in the proportions of the sample population using the various products. There is also a wide range in the fractions of users who report dissatisfaction. The highest rates of reported dissatisfaction for any of the 63 items occurred with products used by comparatively small proportions of the sample; for example, 28.6% of the purchasers of pool tables/card tables reported dissatisfaction (Table 3) and 25.0% of the purchasers of used pickup or panel trucks were dissatisfied (Table 4). The highest rates of dissatisfaction for categories in wide-spread use were related to new and used automobiles (Table 4). The high levels of dissatisfaction with new and used cars confirm conventional complaint statistics. However, product categories with high rates of dissatisfaction among users but with lower proportions of use typically do not show up in complaint statistics.

A previous paper on the services section of the study examined whether the extreme negative responses on the satisfaction scale (quite dissatisfied) are predictive of the frequency of reported dissatisfaction provided at the end of each section of the questionnaire (Day and Bodur, 1978). All Spearman rank correlation coefficients between the two sets of responses were significant beyond the .001 level. In the durables data, the number of "quite dissatisfied" responses shown in the last column of Tables 1-4 and the frequency of reported dissatisfaction in each section were as follows: housing and home furnishings (Table 1),  $r_s = .32$ ; home appliance and personal care products (Table 2),  $r_s = .91$ ; entertainment, recreation and education durables (Table 3),  $r_s = .93$ ; cars and other transportation durables (Table 4),  $r_s = .90$ . Except for the housing and home furnishings section, all correlations were significant beyond the .001 level, reinforcing the earlier results. A speculative reason for the lack of association between scale responses and reported dissatisfaction for housing and home furnishings is that buyers may be dissatisfied with their choices on the basis of taste or visible design features rather than poor quality or substandard performance.

#### Sources of Consumer Dissatisfaction

Consumers who reported high dissatisfaction with a particular durable product were asked to consider an inventory of possible reasons for dissatisfaction and to indicate which reason(s) accounted for their dissatisfaction. If several reasons were checked, the respondents were requested to indicate which one was the most important. Table 5 summarizes the reasons for dissatisfaction given in each of the four sections of durable products.

The most frequently cited reason for dissatisfaction was "The quality of materials was inferior." In percentage terms, this reason ranked highest in all sections except entertainment, recreation and education durables. The percentage of respondents who checked this item ranged from 56.5% for entertainment, recreation and education durables to 75.0% for housing and home furnishings. The second most frequently checked reason for dissatisfaction was "The quality of workmanship was inferior." The percentage of respondents who cited this reason for dissatisfaction ranged from 18.8% for home appliance and

personal care products to 60.9% for entertainment, recreation and education durables.

Table 5 reveals varying patterns in the frequency of other reasons cited for dissatisfaction. For example, problems with warranties were mentioned frequently for cars and other transportation durables but were infrequently mentioned for most other durable products. These results parallel earlier findings with respect to services (Day and Bodur, 1978) to the extent that dissatisfaction with the "quality of materials" and the "quality of workmanship" overshadowed the other possible reasons for dissatisfaction. However, statements about warranties were checked infrequently for services but were checked quite often for durables showing that reasons for dissatisfaction may vary widely over classes of products and services. Table 5 shows that respondents who reported dissatisfaction with durable products tended to be more concerned about product quality issues than with issues related to marketing practices. This was also the case for services (Day and Bodur, 1978)

#### Post-Dissatisfaction Responses

To determine the reasons why many respondents who reported dissatisfaction with various items failed to take any action to resolve their dissatisfaction, subjects were asked to indicate which of four possible reasons best explained why they had taken no action. Table 6 summarizes these responses and shows that about four-fifths of the replies fell into two categories: "I didn't think it was worth the time and effort" and "I didn't think that I could get anyone to do anything about it." These results parallel earlier findings (Gronhaug, 1977) which showed that some consumers seek redress or complain only when they are reasonably confident of obtaining a favorable outcome.

Table 7 summarizes the types of private and public action taken by dissatisfied respondents who sought redress or complained to resolve their dissatisfaction. The total number of private and public actions taken by respondents are shown in the right hand column and a breakdown of the specific responses within the two basic categories is shown in the middle of Table 7. Total actions taken (160) were evenly split between private and public actions. Table 7 shows that 65.0% of the private actions involved decisions to switch brands or warning others about the unsatisfactory consumption experience. Over 80% of public actions involved asking the seller for repairs, replacement or refund or otherwise complaining to the seller. Overall, public actions were evenly split between redress seeking and complaining with the former accounting for 48.8% of the total. A notable finding is that about one-half of the reported actions were of a "private" nature and would not normally be brought to the attention of business firms or consumer protection agencies.

#### SUMMARY AND CONCLUSIONS

This paper has reported some results of analysis of consumer responses to dissatisfaction with durable products using data obtained from a probability sample of 119 households in Bloomington, Indiana. The focus of this paper was on post-evaluation responses and three sets of results were analyzed and discussed: the proportions of highly dissatisfied users, the frequency of reasons for dissatisfaction with various items, and patterns of post-dissatisfaction responses. The results showed widely varying proportions of dissatisfied users over the 63 categories of durable products as well as wide variations in the fraction using particular durables. This suggests that complaint statistics based on simple counts will fail to reveal even exceptionally high rates of dissatisfaction with products which are less widely

used. Reasons for dissatisfaction also varied across the four sections of durable products and concern about product quality or performance tended to overshadow other possible causes of dissatisfaction.

The results reported here parallel and reinforce those reported previously for the services part of the Bloomington study (Day and Bodur, 1978). Analysis of the third section which covers nondurables is nearing completion. Together the three parts will provide a unique set of data spanning the full range of consumer products and services. Hopefully, improvements and extensions of the methodology used in the Bloomington study will soon be applied to a large national sample of consumers to provide a base line for evaluating relative levels of satisfaction as well as a rich data base for researchers on consumer satisfaction, dissatisfaction, and complaining behavior.

#### REFERENCES

Stephen B. Ash, "A Comprehensive Study of Consumer Satisfaction With Durable Products," in H. Keith Hunt, ed., Advances in Consumer Research, Vol. V, Association for Consumer Research, 1978, pp. 254-62.

Ralph L. Day and Muzaffer Bodur, "A Comprehensive Study of Satisfaction With Consumer Services," in Ralph L. Day, ed., Consumer Satisfaction, Dissatisfaction, and Complaining Behavior, Division of Research, Indiana University, 1977, pp. 64-74.

Ralph L. Day and Muzaffer Bodur, "Consumer Response to Dissatisfaction With Services and Intangible Products," in H. Keith Hunt, ed., Advances in Consumer Research, Vol. V., Association for Consumer Research, 1978, pp. 263-72.

Ralph L. Day and E. Laird Landon, Jr., "Collecting Comprehensive Consumer Complaint Data by Survey Research," in Beverlee B. Anderson, ed., Advances for Consumer Research, Vol. III, Association for Consumer Research, 1976, pp. 263-68.

Ralph L. Day and E. Laird Landon, Jr., "Toward A Theory of Consumer Complaining Behaviour," in Arch Woodside, Jagdish Sheth and Peter Bennett, eds., Foundations of Consumer and Industrial Buying Behaviour, American Elsevier, 1977, pp. 425-37.

Kjell Gronhaug, "Exploring Consumer Complaining Behaviours: A Model and Some Empirical Results," in William D. Perreault, ed., Advances in Consumer Research, Vol. IV, Association for Consumer Research, 1977, pp. 159-63.

E. Laird Landon, Jr., "A Model of Consumer Complaint Behavior," in Ralph L. Day, ed., Consumer Satisfaction, Dissatisfaction, and Complaining Behavior, Division of Research, Indiana University, 1977, pp. 31-35.

J. P. Liefeld, F. C. H. Edgecombe and L. Wolfe, "Demographic Characteristics of Canadian Consumer Complaints," Journal of Consumer Affairs, 9 (Summer 1975), 73-80.

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, 7 (Winter 1973), 121-27.

John A. Miller, "Data Reduction Techniques and the Exploration of Satisfaction Segments," in Ralph L. Day, ed., Consumer Satisfaction, Dissatisfaction, and Complaining Behavior, Division of Research, Indiana University, 1977, pp. 103-14.

William R. Thomas and F. Kelly Shuptrine, "The Consumer Complaint Process: Communication and Resolution," Business and Economic Review, 21 (June 1975), 13-22.

Marjorie Wall, Lois E. Dickey and W. Wayne Talarzyk, "Predicting and Profiling Consumer Satisfaction and Propensity to Complain," in Ralph L. Day, ed., Consumer Satisfaction, Dissatisfaction, and Complaining Behavior, Division of Research, Indiana University, 1977, pp. 91-101.

Rex H. Warland, Robert O. Hermann and Jane Willits, "Dissatisfied Consumers: Who Gets Upset and Who Takes What Action," Journal of Consumer Affairs, (Winter 1975), 148-63.

Robert A. Westbrook, "Correlates of Post Purchase Satisfaction With Major Household Appliances," in Ralph L. Day, ed., Consumer Satisfaction, Dissatisfaction, and Complaining Behavior, Division of Research, Indiana University, 1977, pp. 85-90.

Judy Zaichkowsky and John Liefeld, "Personality Profiles of Consumer Complaint Letter Writers," in Ralph L. Day, ed., Consumer Satisfaction, Dissatisfaction and Complaining Behavior, Division of Research, Indiana University, 1977, pp. 124-29.



TABLE 1

## USE AND DISSATISFACTION: HOUSING AND HOME FURNISHINGS

Category of Housing and Home Furnishings	Purchasers of this Durable Category		Users Reporting Dissatisfaction		Users "Quite Dissatisfied"*	
	No.	%	No.	%	No.	%
Single Family House	29	25.0	1	3.4	1	3.4
Condominium or Coop Unit	7	6.0	0	0	1	14.3
Mobile Home	6	5.2	0	0	1	16.7
Vacation Home	5	4.3	0	0	1	20.0
Heating Installation	11	9.5	2	18.2	1	9.1
Additions to Home	18	15.7	1	5.6	1	5.6
Swimming Pool	2	1.7	0	0	1	50.0
"Do It Yourself" Projects	50	42.4	0	0	1	2.0
Purchase of Land	10	8.5	0	0	1	10.0
Draperies, Curtains, Linens, etc.	103	88.0	1	0.9	4	3.9
Electric Blankets, Heating Pads	31	26.3	0	0	2	6.5
Baby Furniture	7	5.9	1	14.3	1	14.3
Sofas, Chairs, Chaise Lounges, etc.	62	52.5	2	3.2	1	1.6
Dining Room, Kitchen Furniture	39	33.1	1	2.6	2	5.1
Bedroom Furniture, Mattresses, etc.	48	41.0	0	0	1	2.1
Den, Porch, Patio and Garden Furniture	22	19.0	1	4.5	3	13.6
Floor Coverings	48	41.0	0	0	2	4.2
Housewares, Tableware, Cooking Utensils	84	71.2	2	2.4	3	3.6
Lamps, Clocks, Bric a Bric, etc.	74	62.7	3	4.1	2	2.7

\*Data in this column is from the item-by-item satisfaction scale and shows the number of purchasers of the product who checked "quite dissatisfied." The Spearman rank correlation between the instances of reported dissatisfaction and the number of "quite dissatisfied" responses is .32, significant at the .10 level.

N = 119. The base for percentages in the number of mentions which is occasionally less than 119.

TABLE 2

## USE AND DISSATISFACTION: HOME APPLIANCE AND PERSONAL CARE PRODUCTS

Category of Home Appliance and Personal Care Products	Purchasers of this Durable Category		Users Reporting Dissatisfaction		Users "Quite Dissatisfied"*	
	No.	%	No.	%	No.	%
Refrigerators, Freezers	34	28.6	3	8.8	3	8.8
Air Conditioners, Dehumidifiers	22	18.8	0	0	0	0
Ranges, Ovens, Grills	28	23.9	2	7.1	2	7.1
Clothes Washers, Dryers, Etc.	36	30.8	0	0	0	0
Vacuum Cleaners, Carpet Sweepers	50	42.4	2	4.0	5	10.0
Sewing Machines, Electric Scissors	24	20.3	0	0	0	0
Garbage Disposers, Trash Compactors	17	14.4	1	5.9	1	5.9
Small Kitchen Appliances	84	71.8	5	5.9	4	4.8
Electric Razors, Electric Hair Clippers	23	19.8	0	0	2	8.7
Electric Hair Dryers, Curlers, etc.	73	62.4	8	10.9	9	12.3
Exercise Machines, etc.	11	9.4	2	18.2	2	18.2
Vibrators and Massagers	9	7.8	0	0	0	0
Eyeglasses, Contact Lenses, etc.	75	64.1	5	6.7	5	6.7
Hearing Aids	2	1.7	0	0	0	0
Medical Appliances or Devices	7	6.0	0	0	0	0

\*Data in this column is from the item-by-item satisfaction scale and shows the number of purchasers of the product who checked "quite dissatisfied." The Spearman rank correlation between the instances of reported dissatisfaction and the number of "quite dissatisfied" responses is .91, significant at the .001 level.

N = 119. The base for percentages in the number of mentions which is occasionally less than 119.

TABLE 3

## USE AND DISSATISFACTION: ENTERTAINMENT, RECREATION AND EDUCATION DURABLES

Categories of Entertainment, recreation and Education Durables	Purchasers of this Durable Category		Users Reporting Dissatisfaction		Users "Quite Dissatisfied"*	
	No.	%	No.	%	No.	%
Television Sets of All Types	54	45.8	3	5.6	5	9.3
Radios, Hi Fi, Stereos, etc.	67	56.8	3	4.5	6	9.0
Home Movie Equipment	11	9.3	0	0	0	0
Pianos, Other Musical Instruments	23	19.5	0	0	0	0
Records and Cassette Recordings	74	62.7	1	1.4	2	2.7
Typewriters, Calculators (Personal Use)	68	57.6	2	2.9	5	7.4
Encyclopedias, Books, etc.	65	56.0	2	3.0	5	7.7
Arts and Crafts Equipment	46	39.0	0	0	0	0
Cameras, Photographic Equipment	37	31.6	0	0	1	2.7
Home Workshop Equipment, Tools	39	33.3	0	0	0	0
Adult Sports Equipment	47	40.2	1	2.1	2	4.3
Boats, Motors, Fishing Equipment	22	18.8	1	4.5	1	4.5
Rifles, Shotguns, Hand Guns	9	7.7	0	0	0	0
Tents, Camping Equipment	26	22.2	0	0	0	0
Luggage, Clothes Bags, etc.	52	44.4	3	5.8	6	11.5
Pool Tables, Card Tables, etc.	7	6.0	2	28.6	1	14.3
Children's Sports Equipment	27	23.1	1	3.7	2	7.4
Children's Games, Toys	40	34.2	2	5.0	3	7.5

\*Data in this column is from the item-by-item satisfaction scale and shows the number of purchasers of the product who checked "quite dissatisfied." The Spearman rank correlation between the instances of reported dissatisfaction and the number of "quite dissatisfied" responses is .93, significant at the .001 level.

N = 119. The base for percentages in the number of mentions which is occasionally less than 119.

TABLE 4

## USE AND DISSATISFACTION: CARS AND OTHER TRANSPORTATION DURABLES

Categories of Cars and Other Transportation Durables	Purchasers of this Durable Category		Users Reporting Dissatisfaction		Users "Quite Dissatisfied"*	
	No.	%	No.	%	No.	%
New Car Purchase	35	30.2	6	17.1	5	14.3
New Pickup/Panel Truck Purchase	5	4.3	0	0	0	0
Used Car Purchase	42	36.2	7	16.7	7	16.7
Used Pickup/Panel Truck Purchase	4	3.5	1	25.0	1	25.0
Car or Truck Lease, Rental	19	16.4	1	5.3	1	5.3
Bicycle Purchase	48	41.4	2	4.2	7	14.6
Motorcycle Purchase	12	10.3	1	8.3	0	0
Travel Home Purchase	1	0.9	0	0	0	0
Airplane Purchase	0	0	0	0	0	0
Tires, Batteries and Accessories	94	79.7	6	6.4	4	4.3
Parts for Home Auto Repair	49	43.4	0	0	0	0

\*Data in this column is from the item-by-item satisfaction scale and shows the number of purchasers of the product who checked "quite dissatisfied." The Spearman rank correlation between the instances of reported dissatisfaction and the number of "quite dissatisfied" responses is .90, significant at the .001 level.

N = 119. The base for percentages in the number of mentions which is occasionally less than 119.

TABLE 5  
REASONS FOR DISSATISFACTION

Reasons for Being Dissatisfied	HOUSING AND HOME FURNISHINGS		HOME APPLIANCE AND PERSONAL CARE		ENTERTAINMENT, RE-CREATION DURABLES		CARS AND OTHER TRANSPORTATION	
	Total Number of Mentions	Named Most Important	Total Number of Mentions	Named Most Important	Total Mentions	Most Important	Total Mentions	Most Important
The quality of materials was inferior.	75.0% <sup>1</sup>	45.5%	59.4% <sup>2</sup>	33.3%	56.5% <sup>3</sup>	35.3%	62.5% <sup>4</sup>	21.7%
The quality of workmanship was inferior.	56.3	9.1	18.8	12.5	60.9	35.3	50.0	13.0
The product had drawbacks that I was not told about when I bought it.	31.3	9.1	28.1	8.3	26.1	0	41.7	21.7
The cost of using the product is higher than I was led to believe.	0	0	3.1	0	0	0	37.5	0
The item that was delivered was different than the one I bought.	0	0	3.1	0	0	0	0	0
The product was damaged when delivered.	18.8	9.1	9.4	8.3	21.7	5.9	16.7	0
I had to wait a long time before the product was delivered.	12.5	9.1	3.1	4.2	4.3	0	16.7	0
The product was misrepresented to me by the salesman.	0	0	6.3	0	13.0	5.9	25.0	8.7
The product was misrepresented in advertisements.	6.3	0	15.6	0	13.0	0	0	0
The credit terms were misrepresented to me.	0	0	0	0	0	0	4.2	0
The warranty did not cover all of the things that went wrong.	18.8	9.1	12.5	0	17.4	5.9	45.8	8.7
Repairs or services under the warranty were unsatisfactory.	18.8	0	25.0	16.7	13.0	0	25.0	13.0
The dealer misrepresented his ability to provide parts and service for the product.	6.3	0	12.5	0	13.0	0	20.8	4.3
I was tricked by a salesman into buying a more expensive model than I needed.	0	0	0	0	4.3	0	0	0
The price that was charged was higher than what I had agreed to pay.	6.3	9.1	0	0	0	0	0	0
The product is unsafe.	6.3	0	21.9	12.5	13.0	11.8	29.2	8.7
The product is bad for the environment.	0	0	0	0	0	0	12.5	0
The product wastes energy resources.	6.3	0	3.1	4.2	0	0	12.5	0
The instructions for using and taking care of the product were incomplete or impossible to read.	6.3	0	3.1	0	4.3	0	4.2	0

<sup>1</sup>Based on 16 instances of reported dissatisfaction involving a total of 31 mentions.  
<sup>2</sup>Based on 32 instances of reported dissatisfaction involving a total of 71 mentions.  
<sup>3</sup>Based on 23 instances of reported dissatisfaction involving a total of 60 mentions.  
<sup>4</sup>Based on a total of 24 instances of reported dissatisfaction involving a total of 97 mentions.

TABLE 6

## REASONS DISSATISFIED RESPONDENTS GAVE FOR TAKING NO ACTIONS

"The one particular reason which best explains why you did not do anything"	Replies	
	No.	%
I didn't think it was worth the time and effort I wanted to do something about it but never got around to it	16	43.2
I didn't think that I could get anyone to do anything about it	2	5.4
I didn't know what to do about it, or who to go to for help	13	35.1
	6	16.3
TOTAL	37	100.0

TABLE 7

## SUMMARY OF ACTIONS TAKEN IN RESPONSE TO DISSATISFACTION

Response	Type of Action	Private Actions:		No. of Actions	
		No.	%		
I. PRIVATE ACTION:				80	50.0%
	Decided not to buy that brand of the product again;	26	32.5		
	Decided to quit using that kind of product;	12	15.0		
	Decided to stop shopping at the store where I bought the product;	16	20.0		
	Warned family and friends about the brand, product or store;	26	32.5		
	TOTAL	80	100.0		
II. PUBLIC ACTION:				80	50.0%
	Asked the seller for repairs, replacement, or refund;	39	48.8		
	Complained to: seller;	26	32.5		
	manufacturer;	10	12.5		
	Better Business Bureau;	2	2.5		
	governmental agency or public official;	0	0		
	private consumer advocate or organization;	1	1.2		
	Contacted a lawyer or otherwise took legal action;	2	2.5		
	TOTAL	80	100.0	160	100.0%

# DISSATISFACTION ATTRIBUTIONS AND CONSUMER COMPLAINT BEHAVIOR<sup>1</sup>

S. Krishnan (student), University of Pittsburgh  
Valerie A. Valle, University of Pittsburgh

## Abstract

An attribution theory framework is used to examine the impact of attributions about responsibility for post purchase dissatisfaction on consumer complaint behavior. An empirical taxonomy of complaint behavior is first developed and resultant types of complaint behavior are examined in the context of a framework of psychological distance of attributions about cause of dissatisfaction. Attributions were found to be an important mediator in the explanation of complaint behavior.

Systematic application of theoretical constructs is necessary for a better understanding of consumer complaint behavior. Attribution theory, which is concerned with the perceived causes of events, holds great promise as a relevant theoretical approach.

## Conceptual Framework

The basic premise of this paper is that attributions of responsibility act as a mediator between a consumer's reaction to a product and the behavioral response that follows. Certainly a consumer who feels dissatisfied because he or she was foolish when making a purchase will react differently than one who feels that the manufacturer is responsible for the dissatisfaction.

## Introduction

The study of consumer complaint behavior is important from both a marketing management and a social welfare perspective. Consumer complaints constitute an important feedback mechanism for marketing management to monitor consumer satisfaction with their products and services. Non-complaint by dissatisfied consumers blocks off this feedback. Such a consumer may decide not to buy the product again or warn family and friends to avoid the product, which can be costly to the marketer. Without knowledge of the sources of dissatisfaction, management will be unable to change its procedures. From a social welfare perspective, it becomes important to study remedial action seeking behavior of disadvantaged consumer groups. Such an understanding is important for formulation of consumer protection legislations as well as plans for making consumer protection programs more effective.

Most of the research relating attributions to consumer behavior has been rather loosely based on the theory developed by Harold Kelley (1967), such as the work by Settle and his associates (Settle 1972; Settle and Golden 1974). The problems associated with this early work have been summarized by Burnkrant (1975). The theoretical base for this paper is the two dimensional schema developed by Weiner et al. (1972), describing the possible attributions following an achievement situation. A success or failure can be attributed either to something about the actor (internally) or to something about the environment or situation (externally). In addition, the performance can be attributed to something which does not vary over time (stable) or to something which varies over time (unstable).

Until recently consumer complaint behavior has been a relatively neglected area of marketing research. The main thrust of research in this field is reflected in Hunt (1977) and Day (1977). This material has been largely restricted to descriptive studies of different types of complaint behavior. These studies include empirical research on incidence of complaining behavior, relating demographic variables to complaint behavior (Granobis, Summers, Frazier 1977; Kraft 1977), analysis of records of consumer protection agencies (Liefeld, Edgecombe, Wolfe 1975), records of companies that have received complaint letters (Diener 1977), broader surveys of consumers (Warland, Herrmann, Willits, 1975; Gronhaug 1977). However, these approaches lack explanatory power and are insufficient to capture the underlying process resulting in complaint behavior.

Attributions made when dissatisfied with a product or service can be fit into this two dimensional schema (see Figure 1). For example, a person who believes he received a bad product because he does not have the ability to deal adequately in the marketplace is making a stable, internal attribution. Someone who feels dissatisfied because she didn't spend enough time shopping is making an internal, unstable attribution. The person who blames a bad product on the nature of the manufacturing company is making a stable, external attribution. Finally, perceiving that a product is a "lemon" (i.e., an unusual example of a poor product) is an unstable, external attribution.

Day & Landon (1977) outline the steps toward the development of a theory of consumer complaint behavior. Though their conceptualization delineates circumstances leading to consumer dissatisfaction, a broad typology of complaint behavior, factors influencing complaint behavior and a preliminary decision model of complaint behavior, no empirical evidence is cited to substantiate their theoretical framework. Landon (1977) has proposed a model of consumer complaint behavior where the major explanatory variables suggested are: dissatisfaction, importance, benefit from complaining, and personality. His formulation, however, remains largely atheoretical.

FIGURE 1

EXAMPLES OF ATTRIBUTIONS FOLLOWING  
DISSATISFACTION WITH A PRODUCT

	Internal	External
Stable	Lack of ability	Nature of manufacturer
Unstable	Insufficient time spent shopping	Luck, "lemon"

For achievement situations, it has been found that whether an attribution is internal or external effects a person's response to the situation. An internal attribution leads to more pride following success and

<sup>1</sup> This study was supported by the Administration on Aging, U.S. Department of Health, Education, and Welfare (Grant No. 5-38131)

shame following failures (Feather, 1967). In addition, an observer is more likely to reward or punish an individual who was perceived to be responsible for his/her own success or failure (Weiner & Kukla, 1970). It is predicted that the type of attribution made about a product performance will affect the type of response made by the consumer.

Valle and Wallendorf (1977) report a pilot study using open ended interviews which investigated whether attributions about a product's performance fell into dimensions similar to those relevant for achievement situations. They found that the attributions could be coded into a locus of control dimension, but that it was more complex than simply internal versus external. They suggested a widened classification labeled "psychological distance from the consumer". At one end is an attribution to oneself (e.g., one's own shopping ability) which corresponds to an internal attribution. The other categories of attributions include (in increasing "externality") people known by the consumer (e.g., friends), the retailer and representatives (e.g., sales clerks or service personnel), the manufacturer, and the larger social system (e.g., laws, all manufacturers). They report that the type of action taken by a dissatisfied consumer tended to be consistent with the psychological distance of the attribution.

In this paper an empirical taxonomy of complaint behavior is first established and then the relationships between types of attribution of blame and types of complaint behavior is examined.

#### Research Background and Methodology

The data used in this study is part of a research effort investigating the consumer problems of the elderly with special focus on susceptibility to unfair marketing practices, information processing, perception of the market place, dissatisfaction and complaint behavior. A series of two mail questionnaires with structured questions were used as the primary data collection instrument. The questionnaires were developed through focus group interviews of senior citizens in a large metropolitan area and its rural suburbs. The first wave of mail survey had 2849 useable responses with 62.2% elderly respondents (i.e., over 65 years of age). The analysis includes only the first wave responses of the elderly consumers.

Respondents answered a series of questions regarding their "bad buying experiences". A bad buying experience was defined as any buying experience where the respondent had felt dissatisfied with the product or service, or had felt that he/she had been cheated or taken advantage of. After describing the product/service involved in the bad buying experience and the nature of dissatisfaction encountered the respondents answered two questions regarding attributions of blame for the bad buying experience as well as the actions they took after the bad buying experience. Table 1 and Table 2 show the list of attributions of blame and complaint actions, respectively, used in these questions

TABLE 1

#### ATTRIBUTIONS FOR DISSATISFACTION

What person or persons were to blame for the problems you had with the product or service? ("X" ALL THAT APPLY).

1. The designer or manufacturers of the product
2. The advertisers of the product/service
3. The company who distributed or sold product/service

4. The salesperson/serviceperson involved
5. My spouse, friend, etc., who was involved in the purchase
6. Me

TABLE 2

#### ACTIONS FOLLOWING DISSATISFACTION

Please "X" below any actions you took because of the problem you had.

1. Stopped payment or refused to pay
2. Decided not to buy that product or service or deal with that company again
3. Complained to family or friends
4. Asked for replacement or refund
5. Complained to the person who sold me the product or service
6. Complained to the company or store
7. Complained to a consumer agency
8. Complained to a public agency or my congressman
9. Complained in a letter to a newspaper or magazine
10. Considered taking legal action
11. Consulted or hired a lawyer to protect my interests
12. I took no action at all

#### Analysis and Discussion

A principal components analysis was performed on responses to items in Table 2 regarding complaint behavior, to uncover the underlying structure of complaint actions. An eigenvalue of 1.0 was used to limit the number of factors extracted. Table 3 shows the varimax rotated factor loadings matrix. The stability of the factor loadings was checked by performing two independent factor analyses on odd and even sequence numbered cases and computing the Harman (1967) reliability coefficient. The reliability coefficient for the two split-halves was 0.86.

#### Factor Interpretation:

Factor 1 has the following characteristics:

- complain to family and friends following the bad buying experience
- decide not to buy the product or service again in future
- complain to salesperson involved in the purchase

This type of complaint behavior fits with Day and Landon (1977) classification of 'Private Action' and constitute actions that are non-assertive in nature. There is no attempt to involve public institutions. Nor is there an attempt to seek direct remedial actions. This factor was accordingly labelled "Private Action Type".

Factor 2 has the following characteristics:

- hire lawyer
- consider legal alternatives to get remedial action
- complain to public agencies, senators, etc.
- stop payment to company involved

This form of complaint action concentrates on legal remedies with appropriate involvement of public agencies to support such legal actions. Accordingly, this factor was labelled "Legal Action Type".

Factor 3 has the following characteristics:

- complain to the company involved

- ask for refund or replacement
- complain to consumer agencies, etc.
- write complaint letters to newspapers

This type of complaint action involved direct confrontation with the selling company as well as 'Public Actions' to facilitate remedial actions. Since the main components of this factor include assertive actions to get remedial action after problem incidence, it was labelled "Remedial Action Seeking Type".

Factor 4 has the following characteristic:

- took no action at all following the bad buying experience

This factor was labelled "Non Complaining Type".

The above typology of complaint behavior has important implications for the study of consumer reactions to post purchase dissatisfaction in general. The influence of dissatisfaction attributions, type of dissatisfaction, consumer demographics on complaint behavior can be examined with respect to this empirical taxonomy.

A one-way analysis of variance was performed with each of the four types of complaint behavior being treated as the dependent variable and each of the attributions of blame (Table 1) as predictors. As per the conceptual framework developed, we can hypothesize that the mean score on Factor 4 should be significantly greater for those who made self attribution of blame when compared to those who did not make such an attribution. Similarly for those making external attributions, the mean on the first 3 Factors would be significantly greater than those who do not make such an attribution. Table 4 contains the results of the one-way analysis of variance.

For Factor 1 (Private Action Type) the mean scores for those who make external type attribution was found to be significantly greater than those who did not. It is interesting to note that in the case of self attribution the mean factor score for those who make such an attribution is significantly lower. Similar differences in means are observed for "Legal Action Type" factor and "Remedial Action Type" factor. In the case of "Non Complaining Type" however, the group mean for those who make self attribution is significantly greater.

TABLE 3

VARIMAX ROTATED FACTOR MATRIX<sup>a</sup>

Variables	Factor 1	Factor 2	Factor 3	Factor 4
1. Stop payment	0.108	[0.206] <sup>b</sup>	-0.021	0.094
2. Not to buy it again	[0.602]	0.069	0.095	-0.049
3. Complain to family, friends	[0.603]	0.063	0.072	-0.063
4. Asked for replacement/refund	0.312	0.003	[0.357]	0.239
5. Complained to salesperson	[0.523]	0.184	0.096	0.274
6. Complained to company involved	0.336	0.016	[0.366]	0.276
7. Complained to consumer protection agency	0.066	0.259	[0.440]	0.016
8. Complained to public agency/congressman	0.067	[0.317]	0.254	0.018
9. Complained in a letter to newspaper/magazine	0.029	0.014	[0.296]	0.021
10. Considered legal action	0.083	[0.490]	0.225	-0.053
11. Hire lawyer	-0.009	[0.524]	-0.002	0.059
12. Took no action at all	0.014	-0.049	-0.048	[0.378]
Explained Variance Per Factor	20.4%	11.5%	9.5%	8.8%
Cumulative	20.4%	31.9%	41.4%	50.2%

a - 1955 cases

b - [ ] indicates the highest loading in each row

TABLE 4

ONE WAY ANALYSIS OF VARIANCE OF FACTOR SCORES FOR VARIOUS ATTRIBUTIONS OF BLAME

Attribution of Blame		Mean Score on Factor 1	Mean Score on Factor 2	Mean Score on Factor 3	Mean Score on Factor 4
Manufacturer	Yes	0.512 <sup>a</sup>	-0.044	0.401 <sup>a</sup>	-0.209 <sup>a</sup>
	No	-0.078	0.007	-0.061	0.032
Advertiser	Yes	0.560 <sup>a</sup>	0.108	0.450 <sup>a</sup>	-0.113
	No	-0.044	-0.008	-0.035	0.009
Company	Yes	0.622 <sup>a</sup>	0.156 <sup>a</sup>	0.382 <sup>a</sup>	-0.166 <sup>a</sup>
	No	-0.194	-0.048	-0.116	0.011
Salesperson/ Serviceperson	Yes	0.4872 <sup>a</sup>	0.145 <sup>a</sup>	0.158 <sup>a</sup>	-0.060
	No	-0.095	-0.028	-0.030	0.012
Spouse/Friend, etc.	Yes	0.322 <sup>a</sup>	0.340 <sup>a</sup>	0.041	0.048
	No	-0.008	-0.013	-0.001	-0.001

Attribution of Blame		Mean Score on Factor 1	Mean Score on Factor 2	Mean Score on Factor 3	Mean Score on Factor 4
Self	Yes	-0.283 <sup>a</sup>	0.035	- .0028	0.347 <sup>a</sup>
	No	0.017	-0.002	0.002	-0.021

Factor 1: Private Action Type  
 Factor 2: Legal Action Type  
 Factor 3: Remedial Action Seeking Type  
 Factor 4: Non Complaining Type

<sup>a</sup>Difference significant at 0.0001 Level

These preliminary results extend directional confirmation that attributions about cause of dissatisfaction act as important mediators in the explanation of complaint behavior. The analysis indicates that external attributions are necessary if any of the first three types of complaint behavior is to result and self attribution of blame, on the other hand, leads to non complaint.

The above analysis has examined only one of the factors, viz., attributions about dissatisfaction in the explanation of complaint behavior. Other factors such as individual consumer attributes, importance of purchase, characteristics of product/service involved, earlier history of bad buying experiences, social integration, awareness of information about consumer protection agencies and consumer rights, etc., would also influence the nature of complaint behavior. The influences of some of these factors is currently being examined with the help of the data base used in this preliminary study.

#### Summary and Conclusions

The preliminary results reported in this paper indicate the importance of the use of an attribution theory framework for explaining consumer complaint behavior on post-purchase dissatisfaction. The empirical taxonomy of complaint behavior presented represents an important classification scheme for understanding and analyses of complaint behavior. Descriptive profiles of consumers (in terms of demographics and personality characteristics) who use one of these modes of complaint behavior as a dominant mode of reaction to post-purchase dissatisfaction would be of considerable interest to consumer protection agencies to design more effective programs and better implement current programs.

The relationship between attributions and complaint behavior has been examined using only the locus of control dimension of Weiner et al., (1972) framework by enlarging the internal-external classification. The other dimensions of stability-instability and intentional-unintentional if operationalized and incorporated could extend the explanatory power that has been achieved with the use of a single dimension framework. Additional research is needed for better operationalization of the 'psychological distance of attribution' and other dimensions of Weiner et al., framework in the development of a process model of consumer complaint behavior.

#### References

- R. E. Burnkrant, "Attribution Theory in Marketing Research: Problems and Prospects," in M.S. Schlinger (ed.), Advances in Consumer Research, Vol. 2, 1975.
- R. L. Day (ed.), Consumer Satisfaction, Dissatisfaction and Complaining Behavior, (Bloomington, Indiana: School of Business, Indiana University, 1977).
- R. L. Day and E. L. Landon, Jr., "Toward a Theory of Consumer Complaining Behavior," in A. G. Woodside, P. D. Bennett and J. N. Sheth (eds.) Foundations of Consumer and Industrial Buying Behavior (New York: American Elsevier, 1977).
- B. J. Diener, "The Fulfillment of Consumer Promotions as a Source of Consumer Dissatisfaction," in H. K. Hunt (ed.) Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction (Cambridge, Massachusetts: Marketing Science Institute, 1977) 333-354.
- N. T. Feather, "Attribution of Responsibility and Valence of Success and Failure in Relation to Initial Confidence and Task Performance," Journal of Personality and Social Psychology, 13, (1969), 129-144.
- D. Granobis, J. O. Summers, and G. L. Frazier, "Correlates of Consumer Expectation and Complaining Behavior," in R. L. Day (ed.) Consumer Satisfaction, Dissatisfaction and Complaining Behavior, (Bloomington Indiana: School of Business, Indiana University, 1977).
- K. Gronhaug, "Exploring Consumer Complaining Behavior: A Model and Some Empirical Results," in W. D. Perrault (ed.) Advances in Consumer Research, Vol. IV, ACR 1977.
- H. H. Harman, "Modern Factor Analysis," (2nd Edition) (Chicago: University of Chicago, 1967).
- H. K. Hunt (ed.) Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction, (Cambridge, Massachusetts: Marketing Science Institute, 1977).
- H. H. Kelley, "Causal Schemata and Attribution Process" in E. E. Jones et. al., (eds.) Attribution: Perceiving the Causes of Behavior, (Morristown, N. J.: General Learning Press, 1972) 151-74.
- F. B. Kraft, "Characteristics of Consumer Complainers and Complaint and Repatronage Behavior," in R. L. Day (ed.) Consumer Satisfaction, Dissatisfaction and Complaining Behavior, (Bloomington, Indiana: School of Business, Indiana University, 1977).
- E. L. Landon, Jr., "A Model of Consumer Complaint Behavior," in R. L. Day (ed.) Consumer Satisfaction, Dissatisfaction and Complaining Behavior, (Bloomington Indiana: School of Business, Indiana University, 1977)
- J. P. Liefeld, F. H. C. Edgecombe and L. Wolfe, "Demographic Characteristics of Canadian Consumer Complainers," The Journal of Consumer Affairs, 7 (Summer, 1975), 73-80.
- R. B. Settle, "Attribution Theory and Acceptance of Information," Journal of Marketing Research, 9 (1972), 85-88.



R. B. Settle and L. L. Golden, "Attribution and Action When a Product Fails," Paper presented at the Association for Consumer Research, Fifth Annual Conference, Chicago, 1974.

V. Valle and M. Wallendorf, "Consumer Attributions of the Cause of Their Product Satisfaction and Dissatisfaction," in R. L. Day (ed.) Consumer Satisfaction, Dissatisfaction and Complaining Behavior, (Bloomington Indiana: School of Business, Indiana University, 1977).

R. H. Warland, R. O. Herrmann and J. Willits, "Dissatisfied Consumers: Who Gets Upset and Who Takes Action," The Journal of Consumer Affairs, 9 (Winter, 1975), 148-163.

B. Weiner, I. Frieze, A. Kukla, L. Reed and R. M. Rosenbaum, "Perceiving the Causes of Success and Failure," in E. E. Jones et al., (eds.) Attribution Perceiving the Causes of Behavior (Morristown, N. J.: General Learning Press, 1972) 95-120.

B. Weiner and A. Kukla, "An Attributional Analysis of Achievement Motivation," Journal of Personality and Social Psychology, 15 (1970) 1-20.

## THREE PAPERS ON CONSUMER SATISFACTION/DISSATISFACTION: A COMMENT

George H. Haines, Jr., University of Toronto

### Abstract

This paper discusses three papers on Consumer Satisfaction/Dissatisfaction. It reviews the three papers contributions to research on this topic. Some comments on implications for future research are presented. The hypothesis is presented that consumer satisfaction/dissatisfaction is an important variable in determining how market operations affect the quality of human life.

### Introduction

The first paper, Consumer Response to Dissatisfaction With Durable Products, presents data on consumer satisfaction developed by the survey method. The data are presented clearly. The second paper, Dissatisfaction Attributions and Consumer Complaint Behavior, uses responses from a mail survey to study whether attributions of responsibility act as a mediator between a consumers' reaction to a product and the behavioral response that follows. The third paper, Conceptual and Methodological Issues in Consumer Satisfaction Research, reviews existing survey and experimental research in consumer satisfaction and proposed that Thibaut and Kelley's (1959) comparison level theory could serve as a theoretical model for consumer satisfaction research.

### Consumer Response To Dissatisfaction With Durable Products

This paper, in a logical sense, follows the preliminary results reported in Ash (1977) and Day and Bodur (1977). These earlier papers discussed in detail the key scales used (see Ash (1977, p. 255)), preliminary results of a comprehensive study of consumer satisfaction with durable products, and consumer response to dissatisfaction with services and intangibles. The present paper gives a clear presentation of empirical data on consumer dissatisfaction with durable products. One disappointment has to be that the data presentation is not clearly and concisely linked to the previous theoretical framework (Day and London, 1977) developed by one of the authors. Explicitly providing such linkage would have aided a reader substantially in getting the entire picture of this research project.

The reported statistics on sample size also deserve comment. I should like to thank Professor Ralph Day for discussing the matter with me. The paper states the initial sample size was 600 dwelling units, and that the overall response rate was 80%. This would mean that 480 questionnaires were returned. However, there were four different questionnaires. The results in this paper are based solely on the durable goods questionnaire, not on the entire 480 returned questionnaires. If one were to suppose that the 480 were divided equally among the four questionnaires, then 120 durable goods questionnaires would have been returned, with 119 of them usable. Actually, this was not quite so, but even so, the first impression one might get from this paper that somehow a very large number of returned questionnaires were not usable is inaccurate. It appears the sampling procedure used generates an acceptably high response rate, and almost all respondents are indeed able to answer the questions asked.

The data in Tables 1-4 show that in many instances re-

ported consumer dissatisfaction is a rare event, somewhat akin, perhaps, to the famous nineteenth century example of cavalry officers being kicked to death by their horses. This is not to say that consumer reports of dissatisfaction are Poisson distributed: in fact, there seems to have been no studies on this. It is to say that these data suggest the hypothesis that for many products consumer dissatisfaction may be poisson distributed. One clear area for future research would be to investigate this hypothesis and its implications from a sampling point of view.

### Dissatisfaction Attributions And Consumer Complaint Behavior

This paper takes an old theory - attribution theory - and employs it in the analysis of complaint behavior. The authors are also careful to frame the study in the beginning with a statement of why consumer complaints should be studied. The authors conclude by stating that "the preliminary results reported in this paper indicate the importance of the use of an attribution theory framework for explaining consumer complaint behavior on post-purchase dissatisfaction". The study presented in this paper shows that there is a relation (both theoretical and empirical) between attributions about dissatisfaction and complaining behavior. Nothing in this paper precludes complaining behavior influencing attributions about dissatisfaction, and the authors have provided no test for causality. It is also noteworthy that, unlike the four point satisfaction-dissatisfaction scale used in the Day papers, the data used here appear to be zero-one, which exacerbates the usual problems of associating verbal labels with factors. This paper appears to be an excellent example of the use of factor analysis to generate empirically based hypotheses, both because of the well reported empirical work and the inherent appeal of the conceptual notions employed.

### Conceptual And Methodological Issues In Consumer Satisfaction Research

The third paper really has two parts. The first part is a careful, tightly written review of empirical research designed to "unravel the determinants of consumer satisfaction/dissatisfaction". Toward the end of this section, the authors put forward the proposition that "satisfaction as a construct...may simply be one measure of an attitude". They follow this with the statement that "If satisfaction with a product is essentially the same as attitude toward a product (then)...". Although the role of purchase circumstances was noted earlier, this statement suggests a narrow view of satisfaction. The data in the Day and Ash paper would lead one to suspect that satisfaction with a product is not the same as attitude toward the product. Nor do the data reported in the Krishnan and Valle paper lend support to this hypothesis. As with economists studying market behavior, the effects of channels of distribution are underestimated. The retailer and/or agent the consumer deals with appears to have a great deal to do with consumer satisfaction.

The last part of this paper attempts to resurrect the Thibaut and Kelley book. The authors gloss over rather lightly Thibaut and Kelley's claim that their notion of comparison level refers to interpersonal relationships,

or (at its most general level) tasks (see Thibaut and Kelley, pp. 21-33 and 95-97). They acknowledge that they are going beyond the theory they reference when they extend the comparison level concept of Thibaut and Kelley to satisfaction with products.

The fundamental fact about Thibaut and Kelley's approach, however, is that it is a framework for organizing data and research more than it is a theory. The cost-benefit paradigm of the book is so general that it could never be disproved: any observations one could ever make about human interaction can be explained, after the event, in terms of the framework presented in Thibaut and Kelley. Thus, although the authors are certainly correct in stating that this framework "may provide a.....basis for understanding post-consumption satisfaction/dissatisfaction with products and services", they have failed to provide a convincing argument that their framework represents a theory. A theory must be capable of being disproved (in some sense). The authors have certainly set a major challenge for themselves in claiming it is possible to convert this framework, originally primarily designed to organize the research on social psychology of groups into a coherent pattern, into a theory of consumer satisfaction. The last section of this paper, however, presents no clear proof that they have met this challenge. One suspects that by the time they have produced such a theory they will have moved, by necessity, rather far away from the Thibaut and Kelley framework. If they are to have a theory instead of a framework, then this is what will have to happen.

#### Directions For Future Research

The central issue these papers raise is why should we be interested in consumer satisfaction. It is tempting to think that a generation of people who grew up listening to "Satisfaction" by the Rolling Stones have all decided to try to find out why one cannot obtain satisfaction in market activities. However, there seems to be no research I am aware of about the influence of societal problems posed in popular music upon research efforts ten years later. Certainly the issue of "why study satisfaction" is not a new one (Czepiel and Rosenberg, 1978). The classical answer to this question seems to be that presented by Krishnan and Valle: we should study the phenomenon so as to help companies cope with it, and so as to help governments to regulate our lives. However, it is possible to study this phenomenon from a more general point of view. We really seem to be exploring the effects of market operations on the quality of human life. Many people, I am sure, do this from an optimistic position that knowledge about alternative institutions and functions and their effects on satisfaction can aid in the creation of a better world, where better means the design and implementation of market systems which improve the quality of human life. There is no doubt that this is an important research area. The periodic market breakdowns in the United States (and some other countries) where consumers go out and loot and burn the stores in their neighborhoods are all too well known. Their effects are all too obvious. The issue that seems to have received less study than it deserves is what are the longer run effects of consumer dissatisfaction, and can something be done about it? That is, the argument is that research on consumer satisfaction ought also to end in helping consumers, as well as companies, governments, and quasi-governmental organizations, some of which impinge on our lives without having any market to discipline their behavior.

Well, what are the longer run effects of consumer dissatisfaction? This was one issue explored in some past work which should be reviewed here (Alexis, Haines, and Simon, forthcoming).

The data used in this analysis consisted of results from two surveys conducted in Rochester, New York. The first survey, conducted in the summer of 1966, was used to develop a cross-sectional regression equation to explain family expenditures for food. The independent variables were family income, family size and composition, race, and mobility. The second survey, conducted in the summer of 1969, also provided data on all these variables. The data also included information on the family decision process, and the families enjoyment of food shopping. Detailed analysis of these results is presented in Alexis, Haines, and Simon, 1972a.

The data on enjoyment of shopping activities is particularly important in the present context. Respondents were asked "How would you rate your feelings about shopping for food?". The possible answers were Really enjoy, Don't mind, Dislike, Don't know, No answer.

How were these data analyzed? The cross-sectional regression equation based on a sample of 270 respondents in the 1966 survey was used to predict family expenditures on food in the 1969 survey. It was first shown (Alexis, Haines, and Simon, 1972a) that this equation predicted the results more accurately than a naive forecasting scheme. The predictions were then compared to the actual reported expenditures in 1969. Each family was classified as to whether they spent more than would be predicted (or less) on their weekly food shopping. Also, each family was classified as to whether their weekly expenditures were predictable. Ninety-five percent prediction intervals were calculated, and each family unit was classified as to whether its actual expenditures lay within the prediction interval or outside of it. This was the predictability criterion. There were one hundred and twenty-seven respondents used in this analysis.

The analysis model of James Coleman (1964) was then used to examine whether systematic deviations existed in (a) the predictability of family expenditures, and (b) the tendency of a family to be an overspender. Enjoyment of shopping and the nature of the family decision making process were the independent variables. Detailed analysis of the results has been presented in Alexis, Haines, and Simon 1972b.

What was found? A general summary of the results would be that not enjoying shopping tends to make expenditures both unpredictable and larger than average. Consumers who report dissatisfaction with shopping tend to do, in comparative terms, a poor job at it. The results support the hypothesis that consumer dissatisfaction creates a feedback loop which perpetuates consumer dissatisfaction. One could certainly argue that the results presented by Day and Ash would not disagree with this proposition.

The issue is, supposing this hypothesis to be true, what can be done about it? Here is the greatest need for future research in consumer satisfaction/dissatisfaction. We need to know whether, in fact and in general, consumer dissatisfaction alters consumer behavior to ensure future dissatisfaction, and if so, how such a destructive feedback loop could be ameliorated. It is not enough to just study consumer behavior in this area so as to alter the behavior of firms and of governments. We must not forget that consumers are an integral part of the marketplace also.

#### References

Marcus Alexis, George H. Haines, Jr., and Leonard S. Simon, Black Consumer Profiles: Food Purchasing in the Inner City, Ann Arbor: Division of Research, University of Michigan, forthcoming.

Marcus Alexis, George H. Haines, Jr., and Leonard S. Simon, "Consumption Behavior of Prisoners: The Case of the Inner City Shopper", Alan R. Andreasen (ed.), Improving Inner-City Marketing, Chicago: American Marketing Association (1972a) pp. 61-94.

Marcus Alexis, George H. Haines, Jr., and Leonard S. Simon, "Neighborhood Effects, Family Decision Making Patterns, and Consumption Expenditures", in: Fred C. Allvine (ed.), Combined Proceedings, American Marketing Association, 1971 Spring and Fall Conferences, Chicago: American Marketing Association, 1972b.

Stephen B. Ash, "A Comprehensive Study of Consumer Satisfaction With Durable Products", in: H. Keith Hunt (ed.), Advances in Consumer Research Vol. V, Ann Arbor: Association For Consumer Research, 1978, pp. 254-262.

James S. Coleman, Introduction to Mathematical Sociology, New York: Free Press, 1964.

Larry A. Czepiel and Larry J. Rosenberg, "The Study of Consumer Satisfaction: Addressing The 'So What' Question", in: H. Keith Hunt (ed.), Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction, Cambridge, Massachusetts: Marketing Science Institute, 1978, pp. 92-119.

Ralph L. Day and Muzafer Bodur, "Consumer Response to Dissatisfaction With Services and Intangibles", in H. Keith Hunt (ed.), Advances in Consumer Research, Vol. V, Ann Arbor: Association For Consumer Research, 1978, pp. 263-272.

Ralph L. Day and E. Laird Landon, Jr., "Toward A Theory of Consumer Complaint Behavior", in Arch Woodside, Jagdish Sheth, and Peter Bennett, (eds.), Foundations of Consumer and Industrial Buying Behavior, New York: American Eselvier, 1977, pp. 425-37.

John W. Thibaut and Harold H. Kelley, The Social Psychology of Groups, New York: John Wiley and Sons, 1959.

## CONSUMER SATISFACTION/DISSATISFACTION: AN OUTSIDER'S VIEW

J. Edward Russo, University of Chicago

The study of consumer satisfaction/dissatisfaction is a very recent addition to consumer research. It derives its impetus from the consumerism movement of 1965-75. Intellectually it is rooted in the measurement of social indicators, combined with some traditional business-oriented analyses of complaint behavior. Recognition of these origins is not just of historical interest. The field is new enough that these roots still largely define the goals and methodology for research.

The goals of consumer satisfaction/dissatisfaction (CS/D) research have been mainly pragmatic, especially the reduction of consumer dissatisfaction both by governmental policy makers and by marketing managers. This applied goal carries a more "engineering" than theory-based flavor. The methodology of the field, adapted from the social indicators movement, is primarily survey techniques, especially mail and telephone questionnaires.

### Three Overriding Issues

Before discussing the individual papers in this session, it is worth acknowledging three larger issues that I see as dominating the study of consumer satisfaction/dissatisfaction. The first of these is the relative roles of engineering and theory building. Should our goal be limited to "engineering" a solution to an applied problem, or should we strive for theories that explain the causes and processes of CS/D? The second two issues are problems that existing research in consumer satisfaction/dissatisfaction is confronting, so far with only limited success. The first I call overtheorizing, the top heavy construction of large theories on small data bases. The second is the validity of the data base itself, i.e., of the phenomena that are to be explained and that direct policy formation.

### Engineering Versus Theory

If one's goal is to build a theory of CS/D, one must consider at least three aspects: the type of theory, one explaining output or process; the range of phenomena to be accounted for; and the need for, or value of, such a theory. There is a trend toward process theories, e.g., of consumer decisions (Bettman, 1978), of product knowledge (Johnson and Russo, 1978) and of advertising (Harris, 1979). Unlike these latter areas, CS/D has a more pragmatic goal and, thus, is more output oriented. The concern is more with how to increase satisfaction, by both public and commercial policies, than with the perceptual, evaluative, learning, memorial, computational or other processes that combine to generate a rating of CS/D.

The range of phenomena that a general theory of CS/D would need to account for is very wide. The field is necessarily concerned with all goods and services. These exhibit great variation over several important attributes, among them purchase price, product knowledge/familiarity, vendor characteristics and regulatory assurances. The construction of a theory this general is a formidable task.

We should also ask ourselves: Do we really need a comprehensive theory? Is a detailed explanation necessary to solve an applied problem? Sometimes the

answer is yes, with the alchemists' task of turning lead into gold as the prototype. But more often than not significant progress can be achieved with a simpler understanding. Medicine exhibits many such examples, one current one being the successful treatment of many cancers by chemotherapy. My judgment is that a comprehensive process theory of CS/D is neither necessary nor possible. Much applied progress can be accomplished without it, and we do not yet have a data base strong enough to support such a theory (more on this below).

In a more positive vein, the question can be rephrased as: What level of explanation of consumer satisfaction/dissatisfaction should we strive for? Let me suggest a middle ground that satisfies the engineering goals of the field without being antitheoretical. At this stage in the development of the field we should expose the primary determinants of consumer satisfaction/dissatisfaction. That is, we should seek to identify the main variables that control CS/D. Implicit in this is the notion of ecological validity. The most important variables are those that explain the largest proportion of the variance in the phenomena of CS/D.

### Overtheorizing

Overtheorizing refers to too high a ratio of theory to evidence. The scientific heart of any new area of research is the set of relevant phenomena. Admittedly, the collection of appropriate data is costly and subject to numerous threats to validity. However it is the foundation upon which any valid theory must be built. The problem of overtheorizing is exacerbated by the greater professional glamour associated with theory construction. However, a closer look at most theories in this area belies any attribution of glamour. They are more creatively borrowed than created. Furthermore, borrowed theories like borrowed clothes tend to fit poorly. I am not suggesting that we should not clothe our data with the best theories available. I do suggest, however, that the low cost of borrowing and high cost of data collection have led to overtheorizing in CS/D. This problem is naturally most acute for a new field, which has a limited wardrobe of theory and which must still bear the development cost of data collection methods.

### Validity of the Data Base

There are at least two major sources of threats to the validity of this research. First, much of the data derives from survey techniques, with its litany of problems caused by selective returns, interviewer interactions or biases, the weak level of the responses, etc. One pair of problems associated with survey data deserve explicit mention. The first is a tendency to sample only complainers. This is especially true when the data analyzed are from a government agency or the files of a company's complaint department. A related problem is the tendency to sample extreme experiences. Even if a representative sample of subjects is successfully queried, the dissatisfaction experiences that they recall will tend to be more extreme ones.

The oversampling of extreme experiences is confounded by the reliance on memory to recreate and evaluate these experiences. Ample evidence suggests that this reconstruction process is subject to stereotypical biases. D'Andrade (1974) has demonstrated that the rating of behavior traits that are supposedly based only on

observed responses are subject to systematic distortion in memory. For example, although antagonism and cooperation are believed to be very dissimilar traits their actual frequencies of emitted interpersonal responses correlated +.17 in the data that he reports. (Both behaviors are related to the intimacy of the interaction.) However, observers' memory-based ratings of these two behaviors correlated -.50. This negative correlation reflects the expected incompatibility between antagonism and cooperation rather than the actual behavior. In responding to an interview on consumer dissatisfaction there may be a tendency for many of the recalled experiences to be reconstructed from general stereotypes rather than from memory of the actual events.

A second area of threats to validity is associated with social psychological experimentation. Krishnan and Valle's paper surveys many of these very convincingly, and more comprehensive reviews are available. I only want to acknowledge that this alternative to survey methods is not without its problems for the empirical researcher.

#### Discussion of the Papers

The discussion of the three papers in this session has two goals. The first is to identify and evaluate what I believe the authors intended to accomplish. The second is to state what I believe they should have tried to accomplish. The papers will be judged against their own goals rather than mine. However, I believe it is important to recognize failures of goals as well as failures of execution.

#### LaTour and Peat

In their paper LaTour and Peat review the expectation based theories that have been applied to consumer satisfaction/dissatisfaction and present another such theory of their own contrivance. Let us consider the review first, which is generally excellent. LaTour and Peat's initial comments on survey research are perceptive and highly relevant, especially the dangers of "fishing" for demographic relationships and the small percentage of variance usually accounted for by whatever is statistically significant. Their critical review of experimental methodology is also thoughtful and informative. Besides drawing on general principles of science, such as the critical role of experimental realism in CS/D research, their criticisms also include problems from psychology (demand artifacts) and from statistics (lack of independence among statistical tests). Overall, the review is perceptive and to the point. It should be valuable to researchers entering the CS/D field or to those engaged in work in related fields.

LaTour and Peat's theoretical contribution is a generalization of Thibaut and Kelley's (1959) comparison level theory to individual product attributes. Although the authors see comparison level theory as offering several advantages over other expectation based theories, I fail to be convinced by their arguments. The partition of the comparison level by Thibaut and Kelley into personal, social and situational components introduces considerable complication, e.g., of identification and measurement, for an unmonstrated increase in explanatory power. More important, LaTour and Peat's generalization of the comparison level (or expectation level) to individual attributes seems to me to be a small and somewhat obvious theoretical step for which a high price is paid. The number of parameters in their formulation is extremely large, and whatever advantages it may have in generality it loses in parsimony. This paper seems typical of

the danger of overtheorizing, an elaborate theory built over an inadequate evidential base.

The quality of the presentation of this paper is unusually high and deserves special note. It was well written generally but the arguments were especially clearly presented to the reader.

What should have been accomplished. The paper reports no test of the proposed theory. The reader has little idea of the value of Thibaut and Kelley's componential analysis of a comparison level in the CS/D context. Given the combinatorial increase in parameters, an empirical test, even one that rejected the theory, would have assured the reader that the theory is testable, i.e., the measurement of all parameters is feasible.

The failure to test the theory also avoids a salutary confrontation with ecological validity. CS/D with what products and services can most benefit from the three component partition of a comparison level, or are most appropriate for a multiattribute representation? Neither of these issues is addressed, as eventually they must, if the theory is to make a useful contribution to the study and reduction of consumer dissatisfaction.

#### Day and Ash

Day and Ash report a survey of complaints with durable products. Their goal, namely to collect a broad data base and point out areas where consumer satisfaction/dissatisfaction is particularly high, is much closer to engineering than to theory. There is no pretense to theory building or even to relating their collected data to existing theory. The paper must surely be judged on the quality of the data and the descriptive analysis that is presented. Such an evaluation is not particularly easy because the presentation is unusually skeletal.

The data reported here represent a major effort that deserves recognition. Furthermore, they are part of a larger program of research that includes a wider range of products and services and that intends to observe a larger sample of consumers. Because one of the authors' goals is testing and eventual revision of their interviewing procedure, several potential problems should be pointed out. It appears that those informants who report less dissatisfaction are rewarded by finishing this self-administered questionnaire more quickly. This probably leads to a bias toward underreporting of dissatisfaction. A second potential problem arises in their validity check. They report correlations in the nineties (for three of four cases) between two measures of dissatisfaction. However, this close agreement might be an artifact of self consistency. Such an artifact is especially likely because while completing the questionnaire subjects could not be monitored. They could easily check an earlier response in order to be assured that a later one is consistent with it.

As with many surveys their methodology is susceptible to demand artifacts. To respondents who did not complain, they asked the question, "Why didn't you do anything?" Their subjects responded almost exclusively with "It wasn't worth the time and effort" or "I didn't expect it would do any good." These answers are overly rational and considerably more attractive than admitting ignorance or forgetfulness, which were the other two possible replies. Other possible explanations, such as the consumer's own fault, were not even presented as alternatives to the subjects. Because all of these latter responses are less socially acceptable, one suspects that they occur more frequently than reported by Day and Ash's interviewees.

My major complaint with the paper is the descriptive analysis. It falls below even minimal standards for the reporting of such data. For example, although the percentage of dissatisfied responses is reported over a 115 different product categories, not a single mean has been calculated and reported, either in the tables (which are quite extensive) or in the text. When these means are calculated the percentage of complaints reported is quite low (roughly 6%) when compared with other studies. For example, Andreasen and Best (1977) report 20% complaint rates using a different survey technique. This low complaint rate enhances the possibility of a response bias caused by the relative ease of reporting satisfaction rather than dissatisfaction.

They also present results for small sample sizes without recognizing the limitations thereof. For example, they report that 28.6% (2 of 7) of the purchasers of used pickup or panel trucks were dissatisfied (Table 4). In addition to the irritation of dealing with three digits of accuracy for such small sample sizes, the dangers of ignoring the limitations of these sample sizes are quite real. Day and Ash conclude from the above data that "the highest rates of reported dissatisfaction for any of the 63 items occurred with products used by comparatively small proportions of the sample." It is difficult to know whether this result is genuine or merely an artifact of the higher variability associated with small samples. Certainly it should not be reported until a deeper statistical analysis is performed.

Finally, the authors make little effort to link their data to existing theories or other bodies of data. The former is not their intention, but certainly the reader should expect citation of the representative empirical literature. Entire bodies of work are conspicuous by their absence, such as the volume by Hunt (1977). The reader is continually frustrated by the failure of the authors to do much more than put the tables of data into print.

What should have been accomplished. As suggested above, the analysis of the data needs considerable elaboration. Presumably Day and Ash are waiting to piece together all parts of their research program. This is all well and good; but at the present this paper is being reported separately, not as one chapter of a larger volume. The authors should have provided more pertinent interpretations of the data reported and appropriate comparisons to other descriptive studies. Certainly such an exercise would be valuable in itself since the authors must eventually do this for their expected monograph.

In fairness to the authors, it must be noted that some of the problems with this paper stem from the format of this conference. Instead of a revision of each paper based on private reviews prior to publication, the paper was not permitted a revision and is now reviewed publicly. The reader would be better served by being able to read a revised version of Day and Ash's paper than by reading these comments.

Krishnan and Valle

Krishnan and Valle help to introduce attribution theory to the consumer satisfaction/dissatisfaction literature. In addition, they present data based on a mail questionnaire survey. Attribution theory offers an important perspective, and they provide relevant supportive data.

The attribution notion is valuable because it highlights a variable that is normally overlooked in CS/D research. In order to understand the link between dissatisfaction and overt complaint behavior it may well be important to consider attributions of responsibility or blame. It might be expected that consumers who blame them-

selves for dissatisfaction with a product are less likely to complain. Indeed, Krishnan and Valle's data confirm this hypothesis.

The experiment has several problems worth noting although some of them are minor. Their report that subjects who attributed responsibility for the dissatisfaction to themselves made fewer overt complaints may be confounded by an artifact. How many subjects are likely to admit that they were to blame yet they complained it was the company's fault. They also are not as careful as they should have been in some of their data presentations. For example they claim that Factor 1 in their analysis is "non-assertive" when one of its components is complaints to a salesperson. The only significance level that they report in their study is .0001. One wonders what interesting relationships, possibly inconsistent with attribution theory, this extremely stringent level is obscuring. Also they tend to overstate their conclusions, e.g., that "external attributions are necessary" if complaint behaviors are to result. I rather suspect, and I know businesspeople would agree, that many consumers will hassle the company in hopes of economic gain even if they themselves are to blame for the product failure. Finally, not all of the statistical methods are as strong as one would like. For example, multivariate analysis of variance might have been used instead of the many univariate analyses of variance that were performed.

Overall, however, the contribution easily outweighs the criticism. Their calling attention to attributions of blame is the appropriate level of theorizing. It does not overtheorize a process model of attribution generation or a multiattribute model of attributions or some other theory more elaborate than the data can bear. In addition to exposing a pertinent factor in controlling CS/D phenomena, Krishnan and Valle report empirical results that test the explanatory value of their construct. The only additional step I would ask is a measurement of the percentage of variance in complaint behavior that is accounted for by attributions of blame. How does this factor compare in importance to such other factors as product/service category, sociodemographics, product knowledge and expectation level? This is a likely question for future research.

#### Acknowledgements

I would like to thank Bill Wilkie for the opportunity to discuss three interesting papers in an area in which I recently knew very little. I hope that my recent and limited familiarity with CS/D research has been sufficient to enable a commentary of value. And I ask the indulgence of the authors whose papers I have just discussed if I have failed to appreciate some merit that a more knowledgeable reviewer might have.

#### References

- James R. Bettman, An Information Processing Theory of Consumer Choice, 1978, forthcoming (draft, August 1977).
- Richard Harris, Information Processing Research in Advertising, 1979, in preparation.
- H. Keith Hunt (ed.), Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction, (Cambridge, Mass: Marketing Science Institute, 1977).
- Eric J. Johnson and J. Edward Russo, "What is Remembered After a Purchase Decision?", Carnegie-Mellon Information Processing Technical Report, November 1978.
- J. W. Thibaut and H. H. Kelley, The Social Psychology of Groups, (New York: John Wiley, 1959).
- R. G. D'Andrade. Memory and the assessment of behavior. In H. M. Blalock, Jr. (Ed.), Measurement in the Social Sciences, Chicago: Aldine, 1974.

George P. Moschis, Georgia State University

### Abstract

Consumer education materials and practices have recently been criticized on the grounds that they teach young people very little about effective consumer behavior. This article presents the results of a large-scale study designed to empirically assess the effectiveness of formal consumer education materials and practices in over one dozen junior high schools and senior high schools offering a wide variety of consumer-related courses. Guidelines for developing economic education materials and programs are suggested.

### Introduction

In recent years, educators have shown increasing interest in consumer socialization, that is, the process by which young people acquire consumption-related skills, knowledge, and attitudes. Because of the consumer education movement and the various public policy issues concerning the effects of promotion on young people, there has been a renewed interest in the education of children for effective interaction with the marketplace. Evidence of this interest is attested to by the growing number of states that have recently included consumer education classes in their school curricula.

The school is usually charged with the responsibility of "preparing the youth to function as adults by giving them the skill, attitude, and knowledge bases necessary for good citizenship and economic self-sufficiency" (Campbell 1969, p. 844). Economic competence, for example, is widely accepted as one of the goals of elementary school education. Among the areas of focus of elementary economic education courses has been the emphasis on knowledge and skills, such as understanding of business terms and practices, some basic vocabulary of economics, intelligent money management, and the ability to select and use goods and services wisely (Gavian and Nanassy 1955).

In spite of the common belief that school is the main source of young people's positive consumer behaviors, the existing consumer education materials and practices have been criticized on the grounds that they teach young people very little about effective consumer behavior (Ward 1974). Recent research findings tend to confirm such criticisms. For example, a national study released by the National Assessment of Educational Progress found that a small percentage of 13 year olds can choose the most economical buy at a supermarket, and few 17 year olds understand basic economic concepts (The National Observer 1977). Two other smaller-scale studies conducted in different states report no relationship between the amount of consumer-related courses the adolescent takes at school and the development of his consumer knowledge (Moore et al 1975 and 1976).

This article presents the findings of a large-scale study of formal consumer education materials and practices and suggests guidelines for developing economic education material and programs. The specific objectives of the study were the following: (1) to assess the effectiveness of consumer-related courses taught at junior and senior high schools; (2) to examine the extent to which adolescent students learn (or do not learn) specific consumer-related skills at school; and (3) to suggest types of consumer skills educators should

attempt to teach to various segments of adolescent students. The specific consumer skills examined in this research were fairly representative of those emphasized in various school consumer-related courses: consumer affairs knowledge, consumer finance management, economic motivations for consumption, information seeking, and effective consumer activities (wise selection and use of products).

### The Study

#### Sample

The sample for this study consisted of 806 adolescents from 13 schools in seven towns and cities in urban, suburban, semi-rural, and rural Wisconsin. Some of the schools were chosen on a convenience basis and some on a random basis. Cooperation was requested from officials at middle schools and senior high schools and questionnaires were delivered to those who agreed to participate. These self-administered questionnaires were filled out by students during regular class sessions and took approximately 30-45 minutes to complete. Most of the classes chosen by school officials to participate in the survey were consumer-related courses such as home economics and consumer education. Because of this, the sample contained a disproportionate number of females, almost two-thirds. The sample was well-balanced, though, with respect to age, geographical location, and social class.

#### Definition and Measurement of Variables

The knowledge of consumer affairs concept comprised two specific variables: knowledge of economic and business concepts and knowledge of consumer-related legislation. Knowledge of economic concepts referred to the accuracy of cognitions held with respect to basic terms in the following areas: economics, banking, finance, insurance, real estate, and marketing. Knowledge of consumer legislation referred to cognitions held with respect to unit pricing, bait advertising, code dating and remedies available to consumers. Respondents were asked to indicate if statements like "When you buy stock, you own part of a company" and "Milk sold in the store must show the last date it can be sold" were true, false, or don't know. The number of correct answers given by each respondent was used as the individual's score. The accuracy index could and did range from 0 to 11; its reliability, as measured by coefficient alpha, was .57.

Consumer finance management referred to the ability to correctly price selected expense items in an average family's monthly budget. Respondents were asked to estimate about how much the average American family with two children and a total monthly income of \$1,000 spends on each of the following items: food, clothes, home expenses, automobile expenses, other expenses, and savings. Respondents were assigned a score of 5 for responses falling approximately within plus or minus 10 percent of the actual expense item estimates, a 4 for responses falling within plus or minus 20 percent of the actual figures, 3 for responses falling within plus or minus 30 percent, a 2 for responses falling within plus or minus 40 percent, and a score of 1 for responses falling approximately within plus or minus 50 percent or more of the actual estimates. The actual estimates for the expense items were obtained from the U.S. Department



of Labor. The accuracy index could range from 6 to 30; its reliability coefficient was .61.

Economic motivations for consumption were operationally defined as cognitive orientations concerning the importance of products' functional and economic features; orientation toward comparison shopping and significant discriminating attributes. This variable was measured on a 0 to 25-point index by summing responses to consumption situations possessing various degrees of such properties. Respondents were asked to check whether they thought it was important to know five different items before buying a bicycle, watch, camera, pocket calculator, or hair dryer:

- Guarantees on different brands,
- Name of company that makes the product,
- Whether any brands are on sale,
- Kinds of materials different brands are made of,
- Quality of store selling a particular brand.

Responses were summed across each item to form a 0 to 5-point index for each item. Thus, scores could range from 0 to 25. The reliability of this five-item scale (coefficient alpha) was .69.

Information seeking was operationally defined as an expressed need to consult various information sources prior to purchase. Measurement of the extent of information seeking was made by summing the number of sources the adolescent might rely upon for information or advice prior to purchasing a camera, hair dryer, pocket calculator, bicycle, or a wrist watch. Alternative information sources were "friends," "TV ads," "salespersons," "consumer reports," "one or both of my parents," and "newspaper or magazine ads." The products were selected on the basis of previous studies, relevance to adolescents' consumer behavior, and amount of socioeconomic and performance risk. Previous researchers have used this approach and suggested its desirability. The alpha reliability coefficient was .37.

Consumer activism refers to the ability to buy and use products and services in a rational and efficient way. It was measured by summing responses to seven items measured on a 5-point, "quite a lot - don't know" scale. Typical items were "I plan how to spend my money," "I carefully read most of the things they write on packages or labels," and "I compare prices and brands before buying something that costs a lot of money." The index could range from 7 to 35; its reliability, as measured by coefficient alpha, was .64.

Formal consumer education referred to the number of consumer-related courses taken at school. Students were asked to state the "Number of courses they have taken" in each of the following areas: consumer education, home economics, economics, environmental sciences, and guidance (job education). They were also asked to write the names of any other courses in which they had studied about consumer matters. Number of courses taken in all areas was summed to form a single index. Finally, socioeconomic status of the adolescent student was measured using Duncan's (1961) socioeconomic scale of occupations.

## Results

### Individual Differences

The first consideration in analyzing data for this study was the extent to which the five consumer-related skills vary by socioeconomic and demographic characteristics of the adolescent student. Previous socialization research suggested that the degree to which young people possess these skills may be affected by maturational factors

(age), social class and sex of the respondent.

Table 1 shows mean values of the five dependent consumer skill measures for younger (6th, 7th, and 8th graders) and older (9th through 12th graders) adolescents. The data show that older adolescents possess the five skills to a significantly greater extent than their younger counterparts, suggesting that maturation may be a factor contributing to the adolescent's development of these skills.

TABLE 1

MEAN VALUES OF DEPENDENT VARIABLES<sup>a</sup>  
FOR YOUNGER AND OLDER ADOLESCENTS<sup>b</sup>

Dependent Variables	Younger Adolescents (N=441)	Older Adolescents (N=365)	t-value
Consumer Knowledge	5.37 (1.82)	6.54 (1.76)	83.70***
Consumer Finance Management	15.41 (4.69)	16.07 (4.92)	4.10*
Economic Motivations	15.56 (5.96)	16.46 (6.02)	4.52*
Information Seeking	11.17 (4.63)	12.07 (5.12)	6.88**
Consumer Activism	23.80 (3.25)	24.56 (3.42)	10.45***

<sup>a</sup>Table entries are mean values for each group. Standard deviations are shown in parentheses.

\*Significant at .05 level.  
\*\*Significant at .01 level.  
\*\*\*Significant at .001 level.

Table 2 shows mean values of the dependent variables for the two main social classes represented in our sample. The table shows that middle class adolescents have greater consumer-related knowledge, they are better able (or prepared) to manage consumer finances of a typical family budget, and they have stronger economic (rational) motivations for consumption than their lower-class counterparts. Middle-class adolescents also scored higher on the information seeking and consumer activism measures than lower-class respondents, but these differences were not statistically significant.

Some significant sex differences also emerged. Table 3 shows that male adolescents have greater consumer knowledge than their female counterparts. Female respondents, however, scored higher on the information seeking measure than male adolescents. No other significant differences emerged.

### Effectiveness of Consumer-Related Courses

Formal consumer education at school was expected to correlate positively with the respondent's consumer knowledge, economic motivations for consumption, information seeking, and consumer activism.

Table 4 shows relationships between each of the five consumer skills and the total number of consumer-related

TABLE 2

MEAN VALUES OF DEPENDENT VARIABLES FOR  
LOWER-CLASS AND MIDDLE-CLASS ADOLESCENTS<sup>a</sup>

Dependent Variables	Lower-Class (N=563)	Middle-Class (N=243)	t-value
Consumer Knowledge	5.79 (1.86)	6.16 (1.92)	6.43**
Consumer Finance Management	15.33 (4.60)	16.60 (4.50)	13.09***
Economic Motivations	15.61 (6.03)	16.81 (5.83)	6.86**
Information Seeking	11.43 (4.87)	11.92 (4.88)	1.70
Consumer Activism	24.06 (3.33)	24.36 (3.38)	1.37

<sup>a</sup>Table entries are mean values for each group. Standard deviations are shown in parentheses.

\*Significant at .05 level.  
\*\*Significant at .01 level.  
\*\*\*Significant at .001 level.

TABLE 3

MEAN VALUES OF DEPENDENT VARIABLES FOR  
MALE AND FEMALE ADOLESCENTS<sup>a</sup>

Dependent Variables	Male Adolescents (N=281)	Female Adolescents (N=525)	t-value
Consumer Knowledge	6.19 (2.00)	5.75 (1.80)	10.16***
Consumer Finance Management	15.45 (4.58)	15.86 (4.62)	1.43
Economic Motivations	15.56 (5.92)	15.18 (6.03)	1.96
Information Seeking	11.11 (5.19)	11.83 (4.68)	4.10*
Consumer Activism	23.94 (3.35)	24.26 (3.34)	1.73

<sup>a</sup>Table entries are mean values for each group. Standard deviations are in parentheses.

\*Significant at .05 level.  
\*\*Significant at .01 level.  
\*\*\*Significant at .001 level.

TABLE 4

RELATIONSHIPS BETWEEN SELECTED CONSUMER  
SKILL MEASURES AND FORMAL CONSUMER EDUCATION,  
CONTROLLING FOR AGE, SOCIAL CLASS, AND SEX<sup>a</sup>

Dependent Variables	Formal Consumer Education	Age	SES	Sex
Consumer Knowledge	.02	.30***	-.00	-.09**
Finance Management	.03	.08*	.11***	.04
Economic Motivations	.04	.04	.09**	.04
Information Seeking	-.01	.07	.07	.06
Consumer Activism	-.04	.11*	.00	.09**

<sup>a</sup>Table entries are partial correlation coefficients (i.e., correlations between each consumer skill and the respective independent variable with the effects of other explanatory factors removed).

\*Significant at .05 level.  
\*\*Significant at .01 level.  
\*\*\*Significant at .001 level.

courses the adolescent had taken at school through Spring 1976, with the effects of the respondent's characteristics (age, socioeconomic background, and sex) removed. The data show that the total consumer-related courses taken at school (the average for the sample was 3.33) do not correlate significantly with any one of the five consumer skills. The correlations are fairly low (0 to  $\pm$  .04) and not statistically significant.

Since the formal consumer education measure was constructed by summing the number of consumer-related courses the respondent had taken at school, the assessment of the effects of formal consumer education may be difficult because different sources may emphasize different aspects of consumer behavior, not necessarily all those that our dependent variables were trying to measure. Thus, it seemed appropriate to analyze the influence of formal consumer education on the development of the five skills in terms of each type of consumer-related course. Furthermore, we considered the possibility that different emphasis may be placed on the various courses taught at junior high schools as compared to those taught at senior high schools. Even the possibility that the child's cognitive development stage may affect his ability to learn the various material was considered.

In examining the effectiveness of each type of consumer-related course on the five consumer skills among younger and older adolescents, no significant differences emerged. The correlations were again small and not statistically significant. These findings suggest that the different content emphasized in the various types of consumer-related courses has no significant effect on the child's learning of any of the five specific skills examined.

This study found little evidence that formal consumer education contributes much to the adolescent's learning of various consumer skills, a finding that is consistent with the results of smaller-scale studies using samples of students from North Dakota and Kentucky. At least three possible factors might account for this finding. First, the instructional material may not contain information useful in teaching the young people the effective aspects of consumption as defined in this research. Second, instructors may be using ineffective methods of teaching socially desirable consumer skills. Third, some other variable is associated with the student's propensity to take consumer-related courses, for example, students who make poor grades at school (are not as capable of learning) might perceive such courses to be of little difficulty. Other researchers have pointed out the first two possibilities.

Future research could examine the reason(s) for failure to learn consumer skills at school. The data in the present study could not answer questions regarding the low correlations found between, for example, the number of consumer-related courses taken at school and the amount of consumer affairs knowledge. Studies in this area could examine the effects of present consumer education materials and practices under carefully controlled experimental conditions.

The results of this study provide some useful guidelines for the development of adult and adolescent consumer education curricula.

First, they suggest the need to reevaluate (a) the content of consumer education materials, (b) practices or methods of teaching economic material to youngsters, (c) goals of consumer education in general, and (d) the targets of students who are to receive certain type(s) of economic education. Some researchers, for example, have suggested that economic education should depart from the traditional pattern of teaching young people things such as descriptions of various institutions and their functions, suggesting instead that it should be based on information processing notions gearing the content of education programs to age-related abilities to comprehend (Ward et al 1975).

Consumer education materials and practices designed to teach adolescents how to be effective consumers should emphasize (1) socially desirable consumer acts (e.g., comparison shopping), (2) economic or rational aspects of the consumer decision making process, (3) use of certain sources of consumer information, (4) knowledge about consumer legal rights and business terms in the marketplace, and (5) skills for budgeting and managing consumer finances.

Consumer educators should also know that the greatest need for consumer education in school appears to exist among adolescents from lower socioeconomic families. Lower-class adolescents should be taught how to budget money, economic concepts, legal rights of consumers in the marketplace, and to develop the economic motivations for consumption possessed to a greater extent by their middle-class counterparts. Furthermore, male adolescents appear to lag behind female adolescents on the ability to use consumer information sources prior to decision making, while female adolescents have a greater need than males to be provided with information that would increase their knowledge about consumer matters.

Earnest Q. Campbell, "Adolescent Socialization." In Handbook of Socialization Theory and Research, Edited by David A. Goslin (Chicago: Rand McNally, 1969).

Otis D. Duncan, "A Socioeconomic Index for All Occupations." In Occupations and Social Status, Edited by Albert J. Reiss, Jr. (New York: Free Press, 1961).

Ruth W. Gavian and Louis C. Nanassy, "Economic Competence as a Goal of Elementary School Education," Elementary School Journal, 55(January, 1955), 270-73.

"How Kids Use What They Learn," The National Observer (April, 1977), p. 6.

Roy L. Moore; George P. Moschis; and Lowndes F. Stephens, "An Exploratory Study of Consumer Role Perceptions in Adolescent Consumer Socialization," Paper presented to the International Communication Association (Chicago, 1975).

Roy L. Moore; Lowndes F. Stephens; and George P. Moschis, "Mass Media and Interpersonal Influence in Adolescent Consumer Socialization," Paper presented to the International Communication Association Conference (Portland, 1976).

Scott L. Ward, "Consumer Socialization," Journal of Consumer Research, 1(September, 1974), 1-14.

Scott Ward; Daniel Wackman; and Ellen Wartella, Children Learning to Buy: The Development of Consumer Information Processing Skills (Cambridge, Mass.: Marketing Science Institute, 1975), Report No. 75-120.

ON CONSUMER DISSATISFACTION: CONSUMER ARBITRATION AS AN  
ALTERNATIVE DISPUTE RESOLUTION MECHANISM

Stephen A. Goodwin, State University of New York at Buffalo  
Vijay Mahajan, Ohio State University  
Bhal J. Bhatt, State University of New York at Buffalo

Abstract

Since its development in 1972, consumer arbitration is emerging as a fast, inexpensive and informal dispute resolution mechanism for both consumers and businesses. Arbitration is viewed as a dispute resolution mechanism which may prove particularly effective in matters involving amounts of money too small to justify use of formal litigation mechanisms. This relatively new and potentially effective redress mechanism is described and evaluated in this paper based on one agency's experience with consumer arbitration.

Introduction

Recent years have witnessed an outpouring of studies and recommendations concerning the causes of and remedies for consumer dissatisfaction. The scope of these studies has included the development of theory of consumer dissatisfaction and complaining behavior (Anderson, 1973; Gronhaug, 1976; Day and Landon, 1977), examinations of the nature of complaints (Sheth and Mammana, 1974; Diamond, 1976; Day and Bodue, 1977; Zaltman, Srivastava and Deshpande, 1977; Day and Landon, 1977; Andreason and Best, 1977), description of the demographics of complainers (Liefeld, Edgecombe and White, 1975; Westbrook, 1976), and evaluations of different dispute resolution mechanisms (Jones and Boyer, 1972; Schrag, 1973; Ittig, 1976).

While disputes between businesses and consumers over quality, performance, warranty and service have occurred since the beginning of mankind's recorded history, the rapid rise in the number and types of such disputes since the mid-1960's has become a subject of concern for business, government, the legal profession and the consumer. It is important to recognize that, while the increasing number of consumer disputes may point to basic problems in our economic system, and certainly require a better understanding of consumer complaining behavior, the settlement of disputes in an expeditious and impartial manner is a major goal advocated by consumer protection groups (Jones and Boyer, 1972). This paper is primarily concerned with a specific dispute resolution mechanism: Consumer Arbitration. This relatively new and potentially effective mechanism will be described and evaluated; one agency's experience with arbitration will be summarized; and implications will be discussed. We are of the opinion that consumer arbitration is generally an unknown quantity (Lippman, 1974; Resnick, 1974; Wexler, 1977); our primary goal, therefore, is to increase the awareness level of marketing educators and practitioners in this regard.

Types and Usage of Action Alternatives  
Available to The Dissatisfied Consumer

Numerous studies have recently been completed which detail the nature and usage of dispute resolution mechanisms. For example, in a telephone survey dealing with complaints recalled by respondents with respect to any type of product or service (Warland, Herman and Willets, 1975), it was found that 25% of the 1,215 respondents in a national probability sample indicated that they had recently experienced extreme dissatisfaction but had done

nothing about it. In another study (Day and Landon, 1977) the response "I didn't do anything at all" was checked by 26.5% of the sample with respect to durable products, and by 16% for services and intangible products. Furthermore, 54% of the respondents indicated they would never again purchase the particular nondurable item that they had designated as the most unsatisfactory and 45% of the respondents reported that they had told their friends about their unsatisfactory experience and urged them to boycott.

Similar results have been found in other studies (Zaltman, Srivastava and Deshpande, 1977; Day and Bodur, 1977) underscoring the fact that many dissatisfied consumers decide not to seek out formal redress mechanisms. In the Day and Bodur study (1977), it was found that only 26.8% of their sample sought any kind of specific redress.

Typical reasons cited for failure to initiate formal complaints and seek specific redress action include "It wasn't worth the time and effort", "I didn't get around to it", "I didn't think it would make any difference", "I didn't know what to do or where to go to get help" (Day and Landon, 1977) and so forth. Of those that do seek specific redress, we find that very few choose some kind of legal action (e.g., 11% in the Zaltman, et al., study; 6% in the Day and Bodur study; 5% in the Day and Landon study). Common reasons for failure to select such legal mechanism alternatives would include cost, inconvenience, ignorance, etc.

Interestingly, consumer arbitration (a quasi-legal alternative) is not even listed as an alternative in the studies discussed above. However, since Better Business Bureaus are the primary developers and users of this form of dispute resolution, we can see how often the BBB's are contacted by complaining consumers. Day and Landon (1977) find that roughly 6% of their sample contacted the BBB. Day and Bodur (1977) find that less than 1% contacted the BBB. Zaltman, et al., (1977) find that roughly 5% of their sample contacted an agency such as the BBB.

These figures underscore the fact that few dissatisfied consumers resort to formal dispute resolution mechanisms, of which consumer arbitration is one alternative. Perhaps one of the biggest reasons is ignorance that such a mechanism exists and what it's all about.

Consumer Arbitration

The National Council of Better Business Bureaus, Inc., has been instrumental in instituting consumer arbitration as a formal, quasi-legal redress mechanism. Such programs were initiated in 1972 and are currently in operation in more than 100 out of 143 local Better Business Bureaus (BBB). Arbitration, as defined by the American Arbitration Association, is the process by which two parties authorize an impartial party or panel to resolve their dispute with a final and binding decision. Arbitration, per se, has developed over the years in a variety of fields (particularly in labor relations, accident claims, and medical malpractice). A thorough review of its history, procedure, etc., can be found in Elkouri and Elkouri (1973).

It is important to point out that Better Business Bureaus

resort to consumer arbitration only if efforts at conciliation or mediation have failed. When complaints are initially made, the BBB tries to resolve the dispute first by conciliation. Here, the BBB merely contacts both parties to the dispute and encourages them to effect a compromise solution and reach an accord without the intervention of a third party. Should this fail, the BBB then assumes the role of a mediator. Mediation may be defined as intervention between conflicting parties to promote reconciliation, settlement or compromise. The key here is once again compromise -- the mediator's function is to persuade the parties to come to a voluntary settlement of their dispute. Neither conciliation nor mediation are legally binding dispute resolution mechanisms. Nevertheless, the BBB seeks to initially solve all complaints/disputes by conciliation or mediation. If both fail, however, then arbitration is suggested as a viable alternative.

Figure 1 highlights the major steps involved in a typical consumer arbitration process.

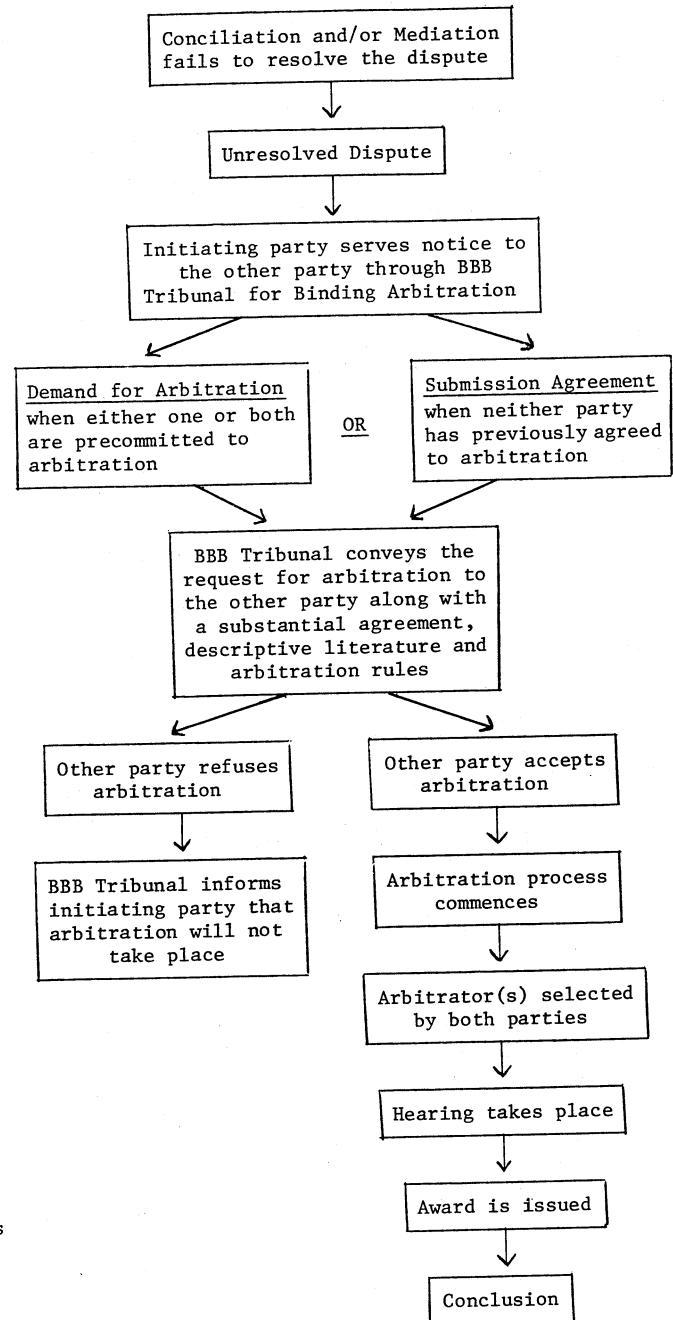
The arbitration process commences when one party serves notice to the other party through the tribunal (BBB) of a desire to submit an existing dispute to an arbitrator for a binding decision. When neither party has previously agreed to arbitration, this notice is in the form of a "submission agreement." If either one or both parties are precommitted to arbitration then this notice is in the form of "demand for arbitration." Precommitment to arbitration normally occurs when a business firm signs an agreement with a "third party", such as the Better Business Bureau, to submit any unresolved dispute to arbitration, or when arbitration is specifically inserted by parties to a contract. The latter case is generally frowned upon. A typical point of view in this regard is that of Rhoda Karpatkin, Executive Director of Consumers Union.

Consumers should not be committed to arbitration at the time of sale by signing such contracts drafted by merchants. They may be deprived of the opportunity to go to court; they may be agreeing to higher fees than courts would charge; they are losing rights to a jury trial or an open hearing, or to have rules of evidence apply; ... and they may be waiving other safeguards enacted to protect debtors and other consumers, and to insure due process in debt collection proceedings. (Wexler, 1977, 41)

The BBB tribunal conveys the request for arbitration to the other party, along with a submission agreement, descriptive literature pertaining to the dispute and a copy of arbitration rules specifying the procedure. If the other party refuses to submit the dispute to arbitration or fails to return the submission agreement within a stipulated time limit, the arbitration process terminates here and the initiating party is so advised. The initiating party may then choose to pursue an alternative redress recourse on his own (e.g., litigation).

Upon receiving a satisfactory execution of the arbitration agreement, the BBB tribunal commences the arbitration process by submitting to both the parties a list of arbitrators with short background profiles of each. With the concurrence of both parties and based on their selection an arbitrator is appointed. Sometimes, a panel of three arbitrators is selected at the suggestion of the tribunal or if stipulated by industry rules or contract provision. In these cases, each party selects one arbitrator from the list provided. These two arbitrators then choose a third arbitrator from the tribunal pool. This third arbitrator serves as chairman of the three person panel. In case of a panel, a majority of two is needed to decide all questions. In case of lack of concurrence among the parties, the agency may appoint arbitrator(s) under special circumstances.

FIGURE 1  
THE CONSUMER ARBITRATION PROCESS



A "Notice of Hearing" is sent to all parties concerned informing of a mutually agreed upon time, place and date. Following the rules of arbitration, an arbitrator hears the case presented by two parties. The parties may be accompanied by attorney if they so desire, but have to inform the tribunal of this decision prior to the hearing. An arbitrator stipulates the issues for arbitration with the concurrence of both parties, takes testimony, evidence, proof, and expert witnesses' testimony and on-sight inspection into account for determining the facts pertaining to the case. Based on the information, an arbitrator writes his "opinion" and issues an "award" recommending remedial redress to the dispute(s). His award is final and binding to both the parties. A challenge to the

arbitrator's award is possible in the courts, but on very narrow grounds in most states. For example in New York State a motion to vacate the award is possible if fraud was involved in procuring the award, partiality of an arbitrator proven, issuing an imperfect award where an arbitrator exceeded his authority, or where either the arbitrator or the administering agency failed to follow provisions of the State statutes.

Upon receiving the arbitrator's award, the tribunal conveys it to both parties informing them of its binding and final nature and requiring its implementation within time limits and conditions stipulated by the award.

#### Advantages of Consumer Arbitration

A thorough review of the sparse literature on consumer arbitration has uncovered a number of dimensions deemed as major advantages of consumer arbitration over other formal dispute resolution mechanisms such as litigation (e.g., in small claims court or full court proceedings, or class action suit) or consumer self-help organizations (e.g., Consumers Education and Protective Association) (National Institute for Consumer Justice, 1972; Wexler, 1971). Advantages include speed, economy, expertise, privacy, informality, convenience, fairness and finality.

According to a National Institute of Consumer Justice report (1972), when comparing consumer arbitration to litigation, the average time to complete the entire arbitration process is roughly 60 days which is substantially shorter than a full court trial dealing with a civil suit.

That arbitration is viewed as being more economical is perhaps a direct result of the "speed" of the process. The arbitration alternative requires minimal expense (For example, in Western New York, the BBB only requires five dollars from the consumer and BBB members; non-BBB members are assessed a charge of thirty dollars. Except under extraordinary conditions, this is the total cost involved to both parties.) to either party. Part of the economy is also derived from the fact that the use of lawyers in the process is minimal; there is no fee for the participating arbitrator(s); and there is generally no fee for expert witnesses.

The expertise dimension pertains to the fact that the arbitrator(s) jointly selected by the parties in dispute is "theoretically" an expert in both arbitration procedures and has pertinent knowledge and background in the sales or product area under consideration. While it is not always possible to assign the "best" arbitrator to each case (in the sense that the arbitrator does have special expertise in the sales or product area involved), the attempt is generally made. In any event, the arbitrator may solicit expert advice when needed.

With regard to privacy, primarily this is advantageous to the business in that arbitration protects business from potentially damaging publicity. The records of the arbitration hearing and award/resolution are not a matter of public record. The media would be hard-pressed to keep track of the proceeding.

As for informality, it is the consumer who benefits primarily because formal court proceedings are likely to be perceived as extremely intimidating. Related to this is the fact that "arbitration allows for independent inquiry and investigation by the arbitrator" (National Institute of Consumer Justice, 1972) thus giving the arbitrator much more flexibility than the courtroom judge. As Wexler (1971) points out, "the arbitrator may submit evidence according to his discretionary test of relevance" and this is advantageous because of "the consumer's limited investigative resources and his ignorance of evidentiary rules."

That convenience is an advantage accrues to the fact that both the time and place of the arbitration hearing can be (and is typically) negotiated by both parties to the dispute. If the consumer can only meet at night or on a weekend, then the business is requested to select a suitable time, and vice versa. The place of the hearing is typically selected so as to be mutually convenient to both parties. There is no cost involved to either party.

The last major advantage, fairness, arises because of the greater degree of flexibility in the arbitration process that "creates an increased likelihood that 'justice' will be done at the initial stage of the dispute resolution process rather than at a later one" (Ittig, 1976, 76). This has special importance in view of the finality of the arbitration decision. A National Institute of Consumer Justice Study (1972) found that "parties by agreeing to arbitrate, voluntarily submit to the final arbitration, and as a result, voluntary compliance hovers around 90%." It is important to reiterate that arbitrated awards can be contested in most states only on the narrow grounds of bias or partiality on the part of the arbitrator. Finally, since the choice of arbitrator(s) is a joint selection, both parties to the dispute have a better chance of having the case heard by an impartial person.

#### Experiences with Consumer Arbitration: A Case Study

Although more than a hundred consumer arbitration programs are known to be in operation throughout the United States, most of these should be considered as test projects (majority of them are less than 5 years old) from which hard facts and tentative conclusions about the process of consumer arbitration are only beginning to emerge. Nationally since the BBB program began 6 years ago, arbitration has been offered in almost 23,000 cases of which 2,300 cases (10%) have resulted in an arbitration hearing. Roughly 6,000 cases (26%) have been settled before the consumer and business agreed to arbitrate, but before the hearing. It has been found that the largest complaint categories are home improvements, car repairs and appliances. The average award value has been established between \$100 and \$200, with a range of \$2.98 recently given to a New Mexico woman to \$14,000 to a Michigan couple (Wexler, 1977). Based on some preliminary data, experiences of one such program offered by the Better Business Bureau in a large northeastern metropolitan area are summarized here (See Tables 1 - 6). Highlights of the preliminary data (being collected since 1973) indicates:

TABLE 1

#### DISPUTES SUBMITTED FOR AWARD THROUGH ARBITRATION

Year	Percentage of Disputes	
	Award Through Arbitration	No Arbitration <sup>1</sup>
1973	16.41	83.59
1974	21.56	78.44
1975	40.80	59.20
1976	38.33	61.67

<sup>1</sup> See Table 2 for reasons.

- (a) In each year arbitration was suggested for between 150 and 200 cases. Notably, roughly one-fourth of these cases were actually resolved through arbitration over the four years examined. However, awards issued through arbitration are on the increase from year to year. In fact, the percentage of cases submitted for award in 1976 (38%) has doubled since 1973 (16%). (See Table 1 and Figure 1.)

(b) As is evident from Table 1, the majority of cases suggested for arbitration are not resolved through this mechanism. The major reasons for non-arbitration are presented in Table 2. Notably, in almost one-third of the cases each year, businesses involved decline to participate in arbitration. On the other hand, in almost one-third of the cases in 1976, consumers either declined or withdrew from arbitration. Finally, over the 4 years previewed, non-arbitrated cases resolved by the BBB tribunal via conciliation or mediation, and/or resolved by mediation during the arbitration hearing, averaged about 18%.

TABLE 2  
REASONS FOR NON-ARBITRATION

Reasons	Percentages of Disputes			
	1973	1974	1975	1976
Not arbitrable	5.02	3.05	5.40	5.40
Business declined	31.05	38.93	39.19	32.43
Consumer declined	1.83	5.34	9.46	13.52
Consumer withdrew request	7.31	5.34	12.16	16.22
Resolved by the tribunal (BBB) through conciliation or mediation	12.79	16.03	16.22	27.03
Resolved by mediation during the arbitration hearing	0.91	1.53	2.70	0.00
Initiating party dropped the matter	41.10	29.78	14.87	5.40

(c) Majority of the disputes were related to home remodeling (57.5%) followed by automotive (16.3%), services (15.6%), retail (18.9%), and financial businesses (0.7%).

TABLE 3  
TYPE OF BUSINESSES INVOLVED IN THE DISPUTE

Year	Percentages of Disputes				
	Auto-motive	Home Remodeling, Const. & Main.	Retail	Services	Financial
1973	6.97	62.80	13.95	16.28	0.00
1974	22.22	47.22	13.89	13.89	2.78
1975	9.81	54.90	7.84	27.45	0.00
1976	26.09	65.21	0.00	8.70	0.00

TABLE 4  
ESTIMATED DOLLAR AMOUNT INVOLVED IN THE DISPUTE

Year	Percentages of Disputes			
	Under \$100	\$100 to \$500	\$501 to \$1000	Over \$1000
1973	39.54	48.84	6.98	4.64
1974	27.78	47.22	13.89	11.11
1975	54.90	33.33	11.77	0.00
1976	21.74	65.22	8.69	4.35

(d) In 80% of the cases each year, the estimated dollar amount in dispute is less than or equal to \$500. Majority of the cases in 1975 and 22% of the cases

in 1976 have estimated values in dispute less than \$100.

(e) In the majority of the cases each year, the award is made in favor of the consumer (54%). In 1976, about 22% were split awards, 26% in favor of businesses and 52% in favor of consumers.

TABLE 5  
TYPE OF RESOLUTION

Year	Percentage of Disputes			
	Split Award	In Favor of Business	In Favor of Customer	Consent Award
1973	13.95	23.26	62.97	0.00
1974	17.14	31.43	48.57	2.86
1975	15.69	31.37	52.94	0.00
1976	21.74	26.09	52.17	0.00

(f) Across the four years, 1973 - 1976, there is a consistent shift towards lesser duration of time involved from the acceptance of arbitration by both parties to the completion of the award. In 1976, about 60% of the cases were completed within 90 days.

TABLE 6  
DURATION OF TIME FOR THE ARBITRATION

Year	Percentage of Disputes				
	Less Than 40 Days	41 to 60 Days	61 to 90 Days	91 to 120 Days	More Than 120 Days
1973	0.00	2.32	9.30	16.28	72.10
1974	0.00	2.78	27.78	16.67	52.77
1975	11.76	17.65	23.53	15.69	31.37
1976	8.69	17.39	34.78	17.39	21.75

In addition, the data also revealed that 9.30% of the cases in 1973, 0% in 1974, 1.96% in 1975, and 4.54% in 1976 used an expert witness. In 4.65% of the cases in 1973, 0% in 1974, 3.92% in 1975, and 4.76% in 1976, either the consumer or the business or both employed a legal counsel.

These preliminary results present the following picture of this program:

Over the four years examined, of the cases suggested for consumer arbitration 25% were resolved through an arbitration hearing. The majority of the cases initiated for arbitration are not submitted for award through arbitration. The major reason for this is the refusal to participate in the arbitration process by either the business or the consumer. However, in each year the percentage of cases resolved through arbitration is increasing. Of all the cases submitted for award, about 60% of the cases were completed within 90 days. The estimated value in dispute in 80% of the cases is less than or equal to \$500. Generally less than 5% of the cases involved a legal counsel (either by the consumer or business or both) or an expert witness. These latter figures strongly support the advantages of the consumer arbitration outlined earlier.

#### Discussion

Consumer arbitration appears to have a great potential

for becoming a satisfactory dispute-resolving mechanism for both consumers and businessmen. Comparing it with other formal redress mechanisms Lippman (1974) writes:

Although the rights of consumers are relatively well established by federal and state enactments and by a plethora of regulatory and remedial agencies, a distinction must necessarily be drawn between having a right and being able to exercise it promptly. Consumer grievances may, of course, be litigated, but not only is such litigation procedurally impractical, it is also prohibitively expensive. A defective twelve dollar toaster or a damaged garment returned from the dry cleaners is simply not grist for our judicial mills as we know them. And, though the class action has been suggested as an answer to the impracticality of litigating individual claims, it is apparent that this highly technical approach is far more appropriate for establishing precedent than for providing the prompt and complete relief the consumer with an everyday marketplace grievance requires.

One attempted solution has been the creation of small claims courts which eliminates papers, bar or discourage attorneys, limit appealability and frequently hold evening sessions -- an important provision lest the litigant barter a day's earnings for recovery which might not even be its equivalent. However, as a solution to minor consumer problems, small claims courts have been largely ineffective, apparently because few aggrieved consumers are aware of their existence ... (a) basic contention here is that there is a remedial group which neither traditional litigation, administrative agencies, class actions, nor small claims courts fill effectively. Informal, inexpensive, speedy, and localized arbitration projects can fill that gap for consumers. (Lippman, 1974, 235-6)

It should be noted that not all complaints can be submitted to arbitration -- for example: instances of clear-cut legal violations; cases where dollar stakes are exceptionally high; or situations where either party wants the right to appear, or desires to set a legal precedent, or seeks damages beyond the sum in dispute.

Some lawyers and consumer protection experts cite other problems with the arbitration process. Since success of the arbitration process depends upon the "willingness of all parties to a dispute to submit themselves to the process ... and to fulfill whatever demands are made upon them as a result" (Ittig, 1976) the opinion has been voiced that many businesses are reluctant to agree to arbitration or have stalled and delayed in arbitration proceedings. Reasons for unwillingness of business to arbitrate range from the knowledge that the "letter of the law is on their side" if not the spirit of the law to a desire to avoid promoting a redress mechanism that is more accessible to consumers and therefore may increase the future number of claims against them (Ittig, 1976).

Since consumer arbitration is but a quasi-judicial redress mechanism, it lacks the force of law to insure the payment of awards issued by the arbitrator.

A number of other issues need to be examined before arbitration can be stamped as a viable alternative. For example, the cornerstone of the consumer arbitration program is the arbitrator. Since cost is a major consideration in consumer arbitration, the arbitrator will, almost of necessity, be a volunteer. The success of this program, therefore, depends to a large extent on the BBB's ability to recruit volunteer arbitrators. It has been proposed that potential arbitrators will come from all walks of life -- retirees, homemakers, ministers,

academicians, and of course lawyers. Questions that need to be investigated are: What type of individuals will make good consumer arbitrators? What type of training programs should be developed to make an ordinary consumer an effective arbitrator (Underhill, 1976)? Will consumer-turned-arbitrator be biased against businesses? Furthermore, although some studies have been conducted (Ittig, 1976) comparing different redress mechanisms, there is a need to empirically evaluate consumer arbitration as compared to other formal redress mechanisms.

Last but not least, the success of this program will depend on its acceptance by consumers and businesses. The data reported in this paper indicate that in almost one-third of the cases initiated the business involved declined to participate in the arbitration and in another one-third (in 1976) of the consumers involved either declined or withdrew from arbitration (Table 1). These figures compiled with the statistics reported earlier from other studies on the low percentage of dissatisfied consumers complaining to BBB's definitely suggest a need for BBB's to bring forth a general awareness of its arbitration program and gain consumers' as well as businesses' confidence as a marketplace neutral. Through the resolution of these issues consumer arbitration may present itself as the dispute resolution mechanism to dissatisfied consumers, businesses, and government.

#### References

- R. E. Anderson, "Consumer Dissatisfaction: The Effect of Disconfirmed Expectancy on Performance," Journal of Marketing Research, 10 (1973), 38-44.
- A. R. Anderson and A. Best, "Consumers Complain - Does Business Respond?" Harvard Business Review, (1977), 104-15.
- R. L. Day and E. L. Landon, Jr., "Toward A Theory of Consumer Complaining Behavior," in A. G. Woodside, J. N. Sheth and P. D. Bennett, eds., Consumer and Industrial Buying Behavior (New York: North-Holland, 1977), 425-37.
- \_\_\_\_\_, and M. Bodur, "Consumer Response to Dissatisfaction with Services and Intangibles," in H. K. Hunt, ed., Advances in Consumer Research, Vol. IV (Chicago: Association for Consumer Research, 1977), 263-72.
- \_\_\_\_\_, and E. L. Landon, Jr., "Collecting Comprehensive Complaint Data by Survey Research," in B. B. Anderson, ed., Advances in Consumer Research, Vol. III (Chicago: Association for Consumer Research, 1976), 263-8.
- S. L. Diamond, S. Ward and R. Faber, "Consumer Problems and Consumerism: Analysis of Calls to A Consumer Hot Line," Journal of Marketing, 40 (1976), 58-62.
- F. Elkouri and A. Elkouri, How Arbitration Works (Washington, D. C.: The Bureau of National Affairs, Inc., 1973).
- K. Gronhaug, "Exploring Consumer Complaining Behavior: A Model and Some Empirical Results," in W. D. Perrault, Jr., ed., Advances in Consumer Research, Vol. IV (Chicago: Association for Consumer Research, 1976), 159-65.
- K. B. Ittig, "The Political Economy of Local Consumer Protection: An Empirical Study of Two Models of Consumer Redress," Ph.D. Dissertation, Department of Consumer Economics and Public Policy, Cornell University, 1976.
- M. G. Jones and B. B. Boyer, "Improving the Duality of Justice in The Marketplace: The Need for Better Consumer Remedies," The George Washington Law Review, 40 (1972), 357-70.



J. P. Liefeld, F. H. C. Edgecombe and L. White, "Demographic Characteristics of Canadian Consumer Complainers," Journal of Consumer Affairs, 9 (1975), 15-35.

M. E. Lippman, "Arbitration as an Alternative to Judicial Settlement," Maine Law Review, (1974), 215-41.

A. N. Resnick, "Consumer Arbitration as an Alternative to Judicial Preseizure Replevin Proceedings," William and Mary Law Review, 16 (1974), 268-93.

P. G. Schrag, Cases and Materials on Consumer Protection (St. Paul: West Publishing Co., 1973).

J. N. Sheth and N. J. Mammana, "Recent Failures in Consumer Protection," California Management Review, 16 (1974), 430-45.

Staff Studies Prepared for The National Institute for Consumer Justice on 1. Business Sponsored Mechanisms for Redress, 2. Arbitration (Washington, D. C.: National Institute for Consumer Justice, 1972).

C. I. Underhill, A Manual for Consumer Arbitrators (Buffalo: Better Business Bureau of Western New York, Inc., 1976).

R. H. Warland, R. O. Herrmann and J. Wilits, "Dissatisfied Consumer: Who Gets Upset and Who Takes Action," Journal of Consumer Affairs, 9 (1976), 148-63.

R. A. Westbrook, "A Study in Consumer Dissatisfaction Before Purchase," in W. D. Perreault, Jr., ed., Advances in Consumer Research, Vol. IV (Chicago: Association for Consumer Research, 1976), 142-8.

R. Wexler, "Consumer Arbitration: A Compromising Position," Juris Doctor, (1977), 40-2.

S. Wexler, "Practicing Law for Poor People," The Yale Law Journal, 6 (1971), 450-65.

G. Zaltman, R. K. Srivastava and R. Deshpande, "Perceptions of Unfair Marketing Practices: Consumerism Implications," in H. K. Hunt, ed., Advances in Consumer Research, Vol. V (Chicago: Association for Consumer Research, 1977), 247-53.

# IN SEARCH OF ACTIONS TO REDUCE CONSUMER SHOPPING PROBLEMS<sup>1</sup>

J. R. Brent Ritchie, University of Calgary  
John D. Claxton, University of British Columbia

## Abstract

The study program reported here has two characteristics of particular interest. First, the use of some relatively new measurement methods, combined with interviewing respondents on three successive occasions, produced an unusually rich data base regarding consumer shopping problems. Second, the study design explicitly included a carefully detailed process for moving from consumer research to action.

## Introduction

Public sector users of consumer research with responsibilities in the areas of consumer assistance/protection are potentially faced with the following complexities:

- (1) understanding what consumers consider to be priority problems.
- (2) assessing the level of government with responsibility for any particular problem.
- (3) determining which existing policies or programs impinge on the problem of concern.
- (4) considering whether industry action can be stimulated or whether government programs have to be developed.
- (5) being sensitive to political pressures that might be initiated by affected interest groups.

Faced by these complexities the difficulty (and importance) of developing action oriented study programs becomes self-evident. The purpose of this paper is to discuss a study program that was instituted to provide guidance in the potential development of consumer assistance/protection programs. This discussion is presented in three major sections. First, there is a brief discussion of the general problem focus of the study. Second, the consumer research that formed the keystone of the study program is described in some detail. Finally, the process used to move from consumer research findings to program development is discussed. The term used to identify this final process is Implications Analysis.

## Focus of Study Program

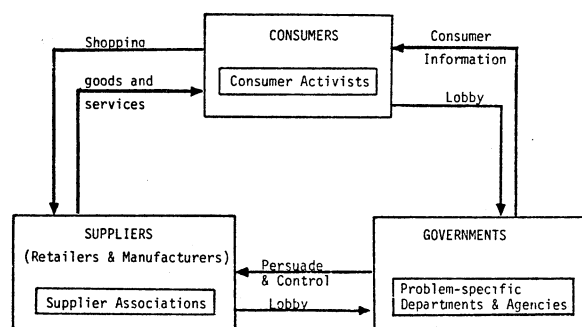
The general focus of the study is provided by the title of this paper, "In Search of Actions to Reduce Consumer Shopping Problems." A framework for studying consumer shopping problems is provided by Figure 1. As indicated in this figure the three major types of participants to be integrated in this type of study are consumers, suppliers, and governments. In the present study these participants were viewed as having the following major roles. Consumers were viewed as being the primary source

of information regarding the definition of shopping problems. Consumers were also used as a source of ideas for solving shopping problems. As a result a consumer-based research study was viewed as the keystone to the study program.

Figure 1 also indicates three interest groups, consumer activists, supplier associations, and problem-specific government departments/agencies. Each of these groups was viewed as a primary participant in the interpretation and application of the consumer research findings. In other words it was these three groups that were identified as the focus for the final stage of the study program, implications analysis.

FIGURE 1

Framework for Studying  
Consumer Shopping Problems



## Purpose of Consumer Research

The general purpose of the consumer study was to identify those product/service areas which pose the greatest difficulties to consumers as they search for and evaluate product/service alternatives. For simplicity these pre-purchase information search and processing problems will be referred to as "shopping problems." The ultimate objective of identifying the most serious problems was to identify public (as well as private) sector actions that might be taken to assist Canadian consumers in overcoming these problems.

For research purposes this overall objective was broken down into a series of four objectives:

1. To identify consumer shopping problems.
2. To evaluate consumer problem priorities by assessing consumer trade-offs when faced with alternative shopping problem situations.
3. To evaluate the reliability/stability of alternative measures of consumer shopping problems.
4. To identify existing and potential actions for reducing consumer shopping problems.

<sup>1</sup>The authors express their appreciation to the Consumer Research Branch of the Department of Consumer and Corporate Affairs, Government of Canada, for the financial support which made possible the research discussed in this paper. They are also grateful to Pierre Filiatrault and Judy Zaichkowsky for their contributions to this study.

The research design consisted of a total of four distinct phases.

Phase I involved development of an annotated bibliography regarding research on prepurchase information search and processing. Further information on this bibliographic work is available from the authors.

Phase II of the research involved data collection and analysis regarding the identification and importance of prepurchase problems (objective 1). This phase began in May 1977 and was completed in August 1977. The methodology used involved Nominal Group Technique (Delbecq, et al., 1975) interviews with consumers in a major urban centre in each of the five regions of Canada. The interviews in each city were equally divided between men and women, and between upper and lower income consumers.

Phase III of the research focused primarily on the trade-offs associated with prepurchase information search and processing (objective 2). The data collection for this phase, which involved extensive use of conjoint measurement (Green and Wind, 1973), began in October 1977 and was completed by November 1977. Data were also gathered concerning consumer perceptions of possible solutions to the problems identified in Phase II. The respondents participating in Phase III were the same as those of Phase II.

Phase IV of the study conducted during the period February - March 1978, involved the collection of the two types of data. The first type was that concerning consumer evaluations of the desirability of a range of previously suggested actions, policies or programs to help them overcome prepurchase information search and processing problems (objective 4). The second type of data gathered was a replication of certain of the conjoint measures employed in Phase III, thus permitting an assessment of measurement reliability/stability (objective 3). Again, respondents for Phase IV were generally the same as those participating in Phases II and III.

#### Measurement - Phase II

The nature of the NGT sessions has been described in detail elsewhere (Claxton, Ritchie and Zaichkowsky, 1978).

These sessions were designed to achieve two goals. First, they sought to identify the major prepurchase information search and processing problems of consumers with respect to each of the five product/service categories. Second, they sought to establish the relative importance of the different problems suggested by the participants in each session.

The structured questionnaire was administered following the NGT sessions. A major thrust of this questionnaire was to explore consumers' beliefs regarding solutions for consumer shopping problems. The discussion of solutions to consumer shopping problems was broken down into two components. First, participants were asked to identify those actions, policies or programs which currently existed and which they found helped them in their shopping or which, in their opinion, made them a more effective or efficient consumer. These ideas were listed on large sheets of paper visible to all participants but no evaluation of the ideas was requested. Subsequently, participants were asked to suggest new actions, policies or programs which they felt might prove useful in assisting consumers to overcome various shopping problems they faced.

#### Measurement - Phase III

Phase III of the study aimed to better understand the trade-offs which consumers would prefer among the most

important problems dimensions identified in Phase II. This trade-off analysis focused on an analysis of shopping problems related to retail services and was restricted to three product/service categories (clothing/footwear, furniture/appliances, and automobile repair services). Four types of data were collected:

- (1) conjoint measures designed to examine consumer trade-offs among the major retail service problem dimensions.
- (2) conjoint measure designed to examine consumer trade-offs among retail service problems and other factors influencing the choice of a particular retailer.
- (3) direct measures of the importance of different retail attribute levels in influencing consumer choice of retail outlets.
- (4) participant suggestions for actions, policies or programs which might be undertaken by the government or by consumers themselves to overcome the major problems of information search and processing evaluated via the conjoint and direct measures.

Conjoint measures related to retail service problem dimensions. The study of trade-offs among information search and processing problems related to retail services employed two parallel forms of conjoint measurement for each of the three product/service categories. The "full profile" or "concept testing" method was implemented using a partial factorial design (orthogonal) requiring the participants to rank order a series of cards containing the profiles of retail outlets in which each of the major problem dimensions was present or not present. Thus, for clothing/footwear and furniture/appliances, four problem dimensions were present at two levels while automobile repair service involved five dimensions at two levels each. The "paired choice" method for measuring consumer trade-offs of problem dimensions requested participants to express their preferences for retail shopping situations in which all possible pairs of shopping problems were presented for consideration. Thus, a total of six dimension pairs were presented to participants in the case of clothing/footwear and furniture/appliances with ten pairs being presented for automobile repair services.

Conjoint measures relating retail service dimensions to other factors influencing choice. In an effort to establish some measure of how important consumers considered the problem dimensions examined above in comparison to other factors influencing the choice of a particular retail outlet, participants were required to evaluate the relative desirability of a number of choice situations. These choice situations involved comparisons along four major factors, each of which was evaluated at either two or three levels within a (2 x 3 x 2 x 3) design. The four major factors were:

- (1) overall reputation of store services: a global measure which attempted to summarize the presence or absence of the four/five retail service problem dimensions examined above.
- (2) travel time: convenience of access to the store.
- (3) shopping time: range of products offered by outlet or nearby (clothing/footwear and furniture/appliances), or speed of service (automobile repair services).
- (4) price levels relative to other retailers.

Direct measures of attribute level importance. These measures were collected by means of a graphic line scale and were intended to serve as a cross check of the results obtained from the conjoint analysis.

Participants suggestions concerning possible solutions. Participants were requested to suggest actions, policies or programs that might be undertaken by government or by consumers themselves to overcome the retail service problems identified in Phase II.

#### Consumer Study: Findings

Although it seemed useful to provide a relatively detailed description of the nature of the consumer study, the purpose here is not to discuss all aspects of the findings. Rather the intention is to provide an example of the study results, and indicate how these results are being utilized in the more general process of developing actions for reducing consumer shopping problems. The findings that are discussed here deal with the problems associated with automobile repair services. This area is presented because of its high priority to a broad cross section of consumers.<sup>1</sup> Readers wishing to review the complete findings or who wish to examine certain results in greater detail are referred to the original study available from the authors.

#### Phase II - NGT Results

Each NGT session concerning automobile repair problems provided two forms of raw data:

- (1) a list of all problems identified by the group. On average, this represented 17 problems per session.
- (2) each participant's rank order of importance for the eight problems he considered the most serious.

The data resulting from the 10 sessions that considered automobile repair services consisted of a total of 170 problem statements. An important characteristic of this data bank of problem statements should be noted. The statements from different sessions were not directly comparable in their raw state since the same issues were not necessarily raised in all NGT sessions, or were not put forward with exactly the same wording. As a result, there existed a need to simplify the structure of the many qualitative statements into a more easily understood form which would permit comparison across NGT sessions. Since this particular data analysis problem was not present in prior planning applications of the NGT approach that employed only a single NGT group, it was necessary to develop a standardized method of analysis which allowed the technique to be extended to multiple NGT session situations. The method developed consisted of four basic steps:

- (1) categorization of initial problem statements into problem themes.
- (2) calculation of a score or index reflecting the importance of each problem theme.
- (3) ranking of problem themes according to their importance index.
- (4) regrouping of problem themes to form major problem dimensions.

Identifying Problem Themes. The purpose of the first step was to look for themes that were common across NGT

<sup>1</sup>The current study indicated that both men and women ranked automobile repairs as the first or second most difficult product/service to shop for. A national study of American consumers indicated that automobile manufacturers and repair garages were the fifth and eighth most important (of 25) priority areas for consumer action (Sentry Insurance, 1977).

sessions. The procedure used was to prepare individual cards for each statement. All statements (from all sessions) were then allocated by judges to different problem categories. Each problem category, or problem theme, involved statements containing essentially the same words or ideas. As might be anticipated, the number of themes established varied with the type of product/service in question (automobile repair services, 24; home repair/renovation services, 22; grocery products, 45; clothing/footwear, 32; and furniture/appliances, 24). The actual themes identified for automobile repair services are presented in Table 1.

TABLE 1

#### PROBLEM THEMES IDENTIFIED FOR AUTOMOBILE REPAIR SERVICES

1. Garage owner's lack of integrity by taking advantage of consumers through unfair practices so they can make more money.
2. Not sure that the time charged represents the real time spent on repairs.
3. Depersonalized service - no contact between owner and mechanic.
4. Not having consumer's interest at heart, lack of professionalism.
5. Lack of guarantee on work.
6. Incomplete servicing/repeat visits required.
7. Difficult to locate competent mechanics or reputable firms.
8. Too many self-serve stations and not enough garages and mechanics to go around.
9. Poor servicing under warranty.
10. Warranty is inadequate/warranty information is misleading.
11. Extra costs arising from warranty work.
12. Estimates hard to obtain or grossly inaccurate.
13. Planned obsolescence/rust.
14. Difficult to do repairs due to metric/electronic design/government standards.
15. Difficult to arrange appointment/servicing of car.
16. Repairs not done when promised.
17. Unavailability and long delivery time for parts.
18. Have to purchase a complete unit when all you need is a small part.
19. Difficult to judge value for money when buying parts.
20. Unbelievable incompetence of mechanics.
21. Work carried out by apprentice but billed at rate of specialist.
22. Wide variation in costs across garages (even oil changes).
23. Cost of repairs much too high.
24. High cost and difficulty in getting foreign cars serviced.

Index of Theme Importance. Step two involved the calculation of an index of theme importance. The index employed was designed to reflect how different measures of importance, namely, the frequency with which a problem was selected by respondents as important and the priority accorded to it when it was selected. Thus, a theme which had been identified in most sessions and which had been ranked highly by most participants would receive a high index score; one selected frequently but considered less important, or selected less frequently but judged important by certain participants, would receive an intermediate index score; and finally a theme selected infrequently and accorded little importance would receive a low index score.

In the present study a participant's score for a particular theme was based on the rank importance accorded by that participant. Statements ranked highest in importance received a score of 8, items second in importance received a score of 7, and so on. Theme scores were then aggregated across all participants so as to provide a summary index of theme importance.

Ranking Themes and Identifying Problem Dimensions. The index of theme importance was used to rank the different problem themes within a product/service category in terms of their approximate order of importance. For many studies, this ranking of themes may represent sufficient analysis since it provides the researcher with a quantitative measure of the importance of the various ideas expressed during the NGT sessions. In the present study, it was desired to further aggregate the problem themes so as to identify what were termed major problem dimensions. These dimensions were subsequently employed as the basis for analyzing consumer retail service trade-offs via conjoint measurement. The manner in which the 24 auto repair service themes (Table 1) were regrouped into problem dimensions provides one specific example of this aggregation process (see Table 2).

While aggregation of themes is useful in certain instances, it should be stressed that it results in a much simplified version of the original data and provides less understanding of the original problem statements as expressed by participants. As such, aggregation of themes into dimensions should not be done with a view to subsequently forgetting their existence since it is at the theme level that the greatest insight into the understanding of consumers appears to be gained.

It is also useful to note the difference between identifying themes and identifying dimensions. The purpose in the former is to aggregate across NGT sessions statements that express essentially the same idea, a process conceptually similar to content analysis (Holsti, 1969; Kassarian, 1977). On the other hand, identification of dimensions is done with a view to providing a structure or framework of themes, a purpose similar to that of taxonomic analysis (Sheath and Sokal, 1973).

#### Phase III - Conjoint Analysis Results

Based on the results of the NGT analysis, a total of seven major problem dimensions were identified for automobile repair service. Six of these were retained for subsequent study via trade-off analysis (general industry conditions were not considered to be a controllable choice criterion for consumers).

The first set of trade-offs participants were asked to make judgments concerning five garage service characteristics derived from the major problem dimensions. Thus, consumers were asked to assess the relative importance of the following factors when selecting an automobile repair service:

(1) the extent to which appointments and estimates are honoured.

- (2) the qualifications (competence) of mechanics.
- (3) the extent to which the garage demonstrates concern for the customer, and his specific desires or requests.
- (4) the strength of the guarantee offered.
- (5) the degree to which the customer is treated fairly and honestly.

TABLE 2

#### MAJOR PROBLEM DIMENSIONS IDENTIFIED FROM PROBLEM THEMES FOR AUTOMOBILE REPAIR SERVICES

Dimension I :	Obtaining time/cost commitments Themes 12, 15 and 16
Dimension II :	Incompetent workmanship Themes 6, 7 and 20
Dimension III:	Problems with guarantees Themes 5, 9, 10 and 11
Dimension IV :	Lack of integrity Themes 1, 2, 13, 18 and 21
Dimension V :	Lack of personal concern for customer Themes 3, 4 and 17
Dimension VI :	Repair cost problems Themes 22, 23 and 24
Dimension VII:	Industry conditions Themes 8, 14 and 19

Table 3 provides the results obtained from the analysis of the conjoint measures collected in regards to these characteristics of the repair service. As seen, the qualifications or competency of the mechanics (1.73) and the integrity of the repair service (1.09) were clearly judged to be the most important dimensions. It should be noted that these results were derived from an aggregate analysis in which the median preference rank for each alternative was employed as the group measure of desirability for that alternative. These median ranks were used as input to the MONANOVA scaling routine (Kruskal, 1965) in accordance with the factorial design employed. This routine provided as output utilities (or importance scores) for each of the five garage characteristics given in Table 3.

These characteristics of repair service were then compared via conjoint measurement with three other choice factors: travel time to garage, service time required, and the cost of the repairs relative to that charged by competing garages. The results of this analysis are also shown in Table 3. Based on this analysis the garage service characteristics and relative cost were clearly the most important choice factors.

#### Phase IV - Consumer Suggested Solution

As indicated earlier the search for solutions to consumer shopping problems was done in two steps. During Phase III data collection respondents were asked in an open-ended manner for their suggested solutions on a problem by problem basis. The problems that were the focus of this questioning were the major problem dimensions that had been developed using NGT. For example, when questioned on auto repairs respondents were asked to suggest potential solutions for each of the problem dimensions indicated in Table 2. These suggestions were then combined via content analysis procedures to form a structured questionnaire for Phase IV data collection.

TABLE 3

## RELATIVE IMPORTANCE (UTILITY) ACCORDED TO REPAIR GARAGE CHARACTERISTICS AND TO OTHER CHOICE FACTORS

## Conjoint Analysis (Aggregate)

First Trade-Off Set:                      Importance (Utility)

## GARAGE SERVICE CHARACTERISTICS

Appointments/Estimates	0.45
Qualified Mechanics	1.73
Concern for Customers	0.35
Guarantee	0.69
Fair Dealings	1.09

Second Trade-Off Set:                      Importance (Utility)

## CHOICE FACTOR

Garage Service Characteristics	1.80
Travel time (to garage)	0.02
Service time	0.43
Relative Cost ( $\pm$ 10 to 20%)	1.25

The list of solutions provided in Table 4 is the set of suggested government actions that could have an impact on auto repair problems. The reader can observe that the solutions could be categorized into four sets.

- (1) solutions requiring more resources for existing policies/programs.
- (2) solutions requiring more laws to control business.
- (3) solutions requiring more education programs.
- (4) solutions requiring more consumer information.

It is also interesting to observe that none of the four most preferred solutions are of the educational or informational (self help) type. All four are regulation based.

## Implications Analysis

The final stages in this study program represented an effort to obtain the maximum impact from the findings of the consumer study. Once data from consumers had been analyzed and the major conclusions identified, a methodology was developed which was designed to identify and evaluate the implications of these conclusions. This methodology, termed implications analysis, involved presenting government officials and private sector managers with the findings within the framework of a systematic procedure designed to determine what they implied with respect to their responsibilities towards consumers and how these responsibilities could best be met.

The approach for this stage of the study was to review with officials from government agencies and supplier associations the major conclusions arising from the consumer study. The objectives of this review were:

- (1) to identify the implications concerning the need for actions, policies and programs which should be undertaken by both the public and private sectors in order to assist consumers,

(2) to establish the relative priorities which should be accorded to the various actions, policies and programs identified above,

(3) to evaluate the means of implementing those actions, policies, and programs judged to have the highest priorities,

(4) to identify the organization, department or agency to which responsibility for implementation of recommended actions should be assigned.

TABLE 4

## CONSUMERS PREFERENCE FOR GOVERNMENT ACTION TO REDUCE AUTO REPAIR PROBLEMS

General Actions (that could impact on auto repairs):	AGREEMENT	
	(Percent of Men)	(Rank)
Expanded complaint bureau facilities	75	8
Increased subsidization of consumer organizations	72	9
Increased penalties for business guilty of unfair practices	93	1
Increased budgets for policing consumer protection laws	79	6
Easier access to small claims courts	86	3
Required warranty and after-sale service	88	2
Agency to control retail prices	34	14
High school consumer education program	80	5
Mass media consumer education program	79	6
Information program re consumer satisfaction/dissatisfaction with retailers	58	11
Shopping center information booths to help consumers	34	14
Auto Repair Specific Actions:		
Licensing garages based on warranties, mechanic qualifications	86	3
Local auto test centers to advise consumers	52	13
Subsidized mechanic training	67	10
Subsidized auto care course for consumers	53	12

## Implications Analysis Process

As indicated earlier the final stage in the study program was designed as a specific attempt to bridge the gap between the acquisition and utilization of knowledge for planning consumer policies and programs. The key steps of the process involved in carrying out this analysis were as follows:

(1) both top and middle level government managers having a potential interest in the findings were identified. These managers were primarily from the Department of Consumer and Corporate Affairs but other interested persons were also involved.

(2) the major conclusions derived from each phase of the research were first presented to a relatively small group of top government officials in an effort to identify major strategic conclusions with respect to the directions which future efforts in consumer protection and assistance should take.

(3) the findings pertaining to each of the five product/service areas studied were presented to middle level government officials having responsibilities for consumer protection with respect to each particular product/service. These managers were asked to review the nature and importance of the problem areas reported by consumers. Opinions were first gathered concerning their perceptions of the completeness of the stated problems and the

consumer's rating of their importance. These middle managers were subsequently asked to identify actions that might be taken by either consumers, industry or government to minimize the most serious problems within each product/service category. This process was followed by a ranking of each potential solution according to its desirability and its feasibility. For those solutions judged both desirable and feasible, specific inputs were sought concerning such issues as:

- (1) who should be primarily responsible for implementation?
- (2) how should the action be financed?
- (3) what time horizon should be envisaged?
- (4) what criteria should be employed to judge the success of the action?

Finally, the results of the discussions with both the top and middle managers were summarized with a view to identifying future action priorities.

#### Conclusions

It is undoubtedly evident to the reader that the implications analysis stage of the research program requires an extended period of time to complete. At the time of writing the implications analysis stage of the study had been conducted with government officials. Still to be completed are the discussions with officials from supplier associations. However, at this point it may be useful to reemphasize characteristics of this research program which are believed to be particularly significant.

First, it is felt that the consumer shopping problems study program represents a major effort at providing consumer policy makers with future oriented research. It is expected that this research can be used to differentiate among real long term needs for consumer protection/assistance, and those which are of a more passing nature.

Second, the particular research methods which were used in the present study were found to be most useful. As such, they may prove to be of interest to other researchers in the field.

Third, it is believed that the emphasis placed on assisting managers to understand and interpret the findings of the research so as to understand their implications for formulating and implementing consumer policies is an essential, although frequently neglected, phase of the study.

#### References

John D. Claxton, J. R. Brent Ritchie, and Judy L. Zaichowsky, "The Nominal Group Technique: Assessing its Potential for Consumer Research," working paper, Faculty of Commerce and Business Administration, University of British Columbia, 1978.

A. L. Delbecq, A. H. Van de Ven, and D. Gustafson, Group Techniques for Program Planning, (Glenview, Illinois: Scott, Foresman and Co., 1975).

P. E. Green, and Y. Wind, Multiattribute Decisions in Marketing, (Hinsdale, Illinois: The Dryden Press, 1973).

O. R. Holsti, Content Analysis for the Social Sciences and Humanities, (Reading, Mass.: Addison-Wesley, 1969).

Harold H. Kassarian, "Content Analysis in Consumer Research," Journal of Consumer Research, 4 (June, 1977), 8-18.

J. B. Kruskal, "Analysis of Factorial Experiments by Estimating Monotone Transformations of the Data," Journal of the Royal Statistical Society, Series B, 27 (1965), 251-63.

Sentry Insurance Co., "Consumers at the Crossroads," conducted by Louis Harris and Associates, and the Marketing Science Institute, 1977.

P. H. A. Sheath, and R. R. Sokal, Numerical Taxonomy, (San Francisco: W. H. Freeman, 1973).

## COMMENTS ON PAPERS ON CONSUMER POLICY PROPOSALS

E. Scott Maynes, Cornell University

In my view successful research provides convincing answers to interesting and/or important questions. Part I of these comments considers the three papers presented in this session as successful or unsuccessful research from this viewpoint. In Part II, I turn to the lessons that these papers hold for future research on consumer policy.

### I

#### Ritchey-Claxton on "Actions to Reduce Consumer Shopping Problems"

The first task of this paper is to uncover an exhaustive list of shopping problems for various products as consumers see them. In a second stage the research seeks to assess the relative importance of these perceived shopping problems. Next, the researchers ascertain consumers' preferences for alternative corrective actions. Finally the major findings of the study are presented, first, to a small group of top government policymakers and, second, to larger groups of middle-level government officials who are asked to evaluate alternative corrective actions as to their desirability and feasibility.

The ACR paper deals only with Stages 1-3.

#### The General Approach

"I shot an arrow into the air ..." This phrase graphically conveys the spirit in which most of us conceive and report our research. Or, more cynically, we craftily design our research with an eye to acceptance by a particular journal (lest we perish!). Not so Professors Ritchie and Claxton. The communication of their findings and the discussion of their implications with policymakers are an integral part of this research plan. For this we can only applaud. Unfortunately, our applause does not extend all of their research design and its execution.

A second feature of this research that does merit our approval is the assumption that parties affected by consumer policy--government officials in particular, but also consumer activists and businessmen--should be familiar with consumer shopping problems as consumers themselves perceive them.

Where I part company from Messrs. Ritchie and Claxton is with their implicit assumption that consumers can pass meaningful judgment on the desirability of various corrective actions. In Table 4, for example, we learn that 75 percent of respondents agree that "expanded complaint bureau" facilities are desirable for dealing with auto repair problems. At what cost to each taxpayer? We learn also that only 34 percent agree that "shopping center information booths to help consumers" are desirable. But what information is assumed to be provided?--(1) platitudinous advice such as "Shopping on credit is more expensive" or (2) relevant local information--for example (an actual example from Minneapolis), the fact that pocket cameras of superior (but equivalent) quality in Consumers Union's judgment can be purchased at prices ranging from \$80 to \$300, and giving the name of the retailer quoting the lower price? Without specification of the information a booth would provide, the recorded attitudes are meaningless.

Thus, I praise Ritchie and Claxton's plan to report and discuss their research findings with appropriate government officials. Certainly, too, public officials as political animals must be aware of public attitudes toward proposed policies, be these attitudes well or ill-informed. But the consideration of appropriate consumer policy requires more than knowledge of consumer attitudes. It requires economic analysis, taking account of prospective benefits and costs.

#### Pilot Study or No: The Nature of the Population and the Sample

Pilot Study or No: The Nature of the Population and the Sample

Research pertaining to an intrinsically uninteresting population (someone's class) or based on a defective sample of an interesting population commands interest only as a pilot study, that is, as an exploration of unknown territory or a demonstration of a new technique or approach.

As consumers of the Ritchie-Claxton paper, you may have already discovered that you are unable to discover the population to which the data in Tables 2-4 allude. Nor will you find any description of the techniques by which this unknown population is sampled. Hence, you will be unable to know whether the fetching data of Table 3 are representative of the respondents only, or of a larger and (perhaps interesting) population. Further research on my part proved inconclusive, but only increased my doubts. In the preliminary version of this paper, the authors wrote:

Participants in the study were obtained through contact with the local headquarters of the Boy Scouts of Canada in each of five cities across Canada. Five cities (Vancouver, Regina, Ottawa, Quebec City, Halifax), were chosen in order to obtain inputs from each of the geographic regions of the country. Within each city the local Boy Scout representatives were asked to provide a total of approximately 60 participants...

Good for the Boy Scouts! But how did they choose? No information for answering this question is provided. Hence, I can only recommend excommunication for Ritchie and Claxton and others like them who fail to provide such essential information.

It would be my judgment that these authors do not break enough new ground to legitimize their research on pilot study grounds.

#### Consumer Shopping Behavior As Terra Incognita

Judging from their references (the Sentry Insurance Company survey excepted), Ritchie and Claxton have approached the problem of consumer shopping problems as though they were exploring virgin wilderness.

We all know that it is not. Presumably reflective of problems encountered in shopping, the first Better Business Bureau was established in 1912! Consumer economics texts (Gordon's Economics for Consumers and Troelstrup's The Consumer in American Society) began discussing consumer shopping problems normatively in the



1930's. A paper of mine (Maynes, 1973) explored the causes of consumer information and shopping problems in 1973. The excellent Best-Andreasen national study of consumer dissatisfaction and complaints was published in 1976 (Best and Andreasen). Finally, two conference proceedings (Day, 1977 and Hunt, 1977) have focused entirely on consumer satisfaction and dissatisfaction, much of it related to shopping.

"Better" research builds on previous knowledge. I chide Msrs. Ritchie and Claxton for approaching mapped territory as though it were totally unexplored.

#### Summing-Up

Comments that began with approval have unfortunately been succeeded by the citation of blemishes that convert this from "successful" to "unsuccessful" research. Its merits are an interesting and useful target, and the feedback of the research findings to officials responsible for formulating and executing corrective policies, a feature highly worthy of evaluation. But apparently all this effort went for naught, lost on a sample that may be representative only of itself. (I would be delighted if information were provided to show that it was a probability sample, satisfactorily executed, or--alternatively--a quota sample that could be shown to be reasonably "representative.") Finally, this research did not build--at least there was no evidence to show it--on past exploration of this area by others--the Better Business Bureaus, the normative consumer economists, and recent consumer behaviorists.

#### Moschis on "An Empirical Assessment of Formal Consumer Education"

When 92 percent of Americans support the introduction of more consumer education in American schools--as the Sentry Insurance survey shows [Sentry, 1976, p. 76], then it becomes important to know whether some variant of formal "consumer education" is effective. Hence, Moschis' research poses an appropriate question. Unfortunately, as the remarks below suggest, his research does not provide a "convincing answer" to the question posed for some reasons that are beyond his control and some that are within his control.

#### The Factors Beyond Moschis' Control

A fundamental problem for this investigation is that no consensus exists as to what constitutes "consumer economics" or "consumer education." As an example, I as the author of an undergraduate text in Consumer Economics would argue that an understanding of price discrimination is a sine qua non for the consumer who would purchase effectively in a contemporary local consumer market.<sup>1/</sup> Price discrimination (and especially the point that it may be initiated by the consumer) is a topic that is emphasized in my text. It is omitted or passed over most cursorily in most texts now on the market.<sup>2/</sup>

<sup>1/</sup> Salop (1977) has modeled local markets in which sellers sort our consumers on the basis of their knowledgeability, charging higher prices to the ignorant and lower prices to the knowledgeable. It would be my judgment that this model is highly realistic.

<sup>2/</sup> Daniel A. McGowan's Consumer Economics, (Chicago: Rand-McNally, 1978) is a welcome exception.

Going further, I would assert that most texts are mainly descriptive ("preachy") while the problems of consumers are analytical: How much should I search? How do local retail markets work? How do I assess the quality of products? Given inflation and taxes, what asset should I choose for the investment of my savings? What's more, the profession has not yet agreed on which disciplinary background--economics, psychology, consumer behavior(!), home economics--should be stressed most strongly. Given this undeveloped, non-consensus state of consumer economics and consumer education at the university level, is it strange that Professor Moschis may encounter difficulties in deciding what consumers should know and be tested on? Or that I should criticize one of his tests as "invalid"?

The test that I single out for criticism is that relating to "consumer financial management." In essence, "mastery" of consumer financial management is signaled by the test when a student-respondent can accurately estimate the percent of income an average American family would spend on such major categories as food, clothes, home expenses, automobile expenses, other expenses, and saving. Though this information may be interesting, I doubt that it aids in effective financial management for "my family." My major doubts are confined to this particular test; the others seem defensible. However, they might be more defensible had they been gleaned from a conference of "experts" in this (diverse) field.

#### The Factors Within Moschis' Control

The Sample. In seeking to test the effectiveness of formal consumer education, Moschis has faced two sampling problems: (1) the selection of consumer education courses or subject matter taught, and (2) the selection of a representative group of students.

Moschis could have selected--and this is what I would have viewed as preferable--a course viewed by some group of "experts" as a "very good" or "model" course. Such a procedure would have the advantage that we would (or could) know the exact contents of the stimulus to which students were exposed. Instead, Moschis chose to select a sample of 13 schools in Wisconsin. In effect, this means that the consumer education materials to which students were exposed are a weighted average of whatever is taught in the "consumer education" courses of these 13 schools.

And there lies the rub. According to Moschis, "some [how many?] of the schools were chosen on a convenience basis and some on a random basis." We have no information on the degree to which either courses or the students enrolled in them are representative of all Wisconsin consumer education courses and students. One can observe that, if school-to-school variation in courses or students is great, a sample of 13 from the hundreds of Wisconsin schools would be frightfully inadequate, even if chosen by probability methods. So I chide Moschis for failing to utilize adequate sampling methods and for failing to provide a detailed description of the methods he did apply. As consumers of this research, we can only hope that this sample is representative! But it would have been appropriate--and is still appropriate--for Moschis to tell the reader in some detail just what consumer education courses students were exposed to, what their content was, and, quantitatively, in what ways students in these courses differed from average students in their schools. We know they were more female; were their other factors operative?

"Significant" Versus "Statistically Significant" Differences. In his analysis Moschis devotes considerable space to asking whether mastery of consumer education materials differs "significantly" by age (younger vs.

older adolescents), social class, and sex. Not very interesting questions. But they provide classic examples of "statistically significant" differences that are not substantively important. For example, from Table 1 we learn that older adolescents score 12.07 (out of 30) on "information seeking, while younger adolescents score only 11.17 (out of 30). A statistically significant difference at the .01 level. But does Moschis contend that a less-than-one-point difference on a rather crude 30 point scale is "important" in determining the effectiveness of formal consumer education?

What interests me far more--and what goes virtually undiscussed--are the mean values of these scales. But of even greater interest is whether students with higher scores have become more effective consumers in real life. Here we have no evidence at all.

#### The Quantification of Formal Consumer Education.

Finally, in Table 4, we come to the key question: does exposure to formal consumer education exert a statistically significant effect (the narrow question first) on the five components of consumer skill--consumer knowledge, finance management, economic motivations, information seeking, consumer activism? The answer of the partial correlation coefficients is emphatically "no." But is this answer convincing, even for this sample? It may not be. The formal education variable, as entered in the equations, was defined as "the number of consumer related courses taken." The mean value was 3.3 courses. My complaint is that each course is treated as exerting a uniform effect; should not the first course taken have been given a greater weight?

Because of the unsatisfactory statistical formulation of the narrow question, there is no point in asking the important question: does exposure to formal consumer education courses contribute to more effective consumer behavior?

#### Summing-Up

For reasons both beyond and within the author's control, this investigation does not, in my judgment, provide a convincing answer to the question posed. But it does provide useful lessons for those to come both with respect to steps to be taken and steps not to be taken. As a profession, we must press on to a convincing answer

Goodwin, Mahajan, and Bhatt on "Consumer Arbitration As An Alternative Dispute Resolution Mechanism"

This is not your usual research paper. It reviews the history and then describes the working of consumer arbitration as a complaint resolution mechanism. Finally data are presented for one "large northeastern metropolitan area" relating to such facets of arbitration as the fraction of cases resulting in an arbitration award, the reasons for non-arbitration, the types of businesses, the dollar amounts, the resolution, and time taken for arbitration.

#### Fraud in Perspective

Following the presentation of sample data on consumer arbitration in one metropolitan area, the authors declare: "Consumer arbitration appears to have a great potential for becoming a satisfactory dispute-resolving mechanism for both consumers and businessmen." In my judgment neither their discussion nor the data they present support this conjecture. Let me say why.

I start with a proposition that a complaint resolution

mechanism that is open only to a tiny fraction (the most persistent and most knowledgeable) complaining consumers is not a "satisfactory" dispute-resolving mechanism even though it treats those who use it fairly. The reason: it may offer no satisfactory avenue for the great mass of consumer complainants. Indeed, if the arbitration mechanism is utilized by a sufficiently small fraction of complainants, it may be appropriate to label it a consumer "fraud" because it gives the appearance of providing justice without providing "satisfactory" complaint resolution to the large majority of complainants. Underlying this observation is the assumption that the time costs and subjective costs (feelings of awkwardness, inadequacy, annoyance, etc.) are likely to be very great on the part of consumer complainants. If they are sufficiently great, consumers may not take the next step in the consumer redress process.

The critical element missing from this paper is the failure to tell us what fraction of grievances presented to the Better Business Bureaus have been taken to arbitration in this sample area. Bear in mind that the Better Business Bureaus process only a small fraction of voiced grievances--0.3 percent of cases where consumers perceived problems and 0.8 percent of voiced consumer complaints (estimated from data published in (Best and Andreasen, 1976, pp. 46-47).

One further doubt on the author's optimism. In my judgment a median resolution time of 68 days (from Table 6) is "speedy" only when compared with litigation.

Summing up, consumer arbitration may turn out to be either (1) a satisfactory dispute-resolving mechanism (it is devoutly to be wished!), (2) a consumer fraud, or (3) something in between. We are in urgent need of further research.

## II

### The Future and Consumer Policy Research

The organizer of these meetings, William Wilkie, asked each discussant to discuss the field in general. It is pleasant indeed to do what comes so naturally!

#### Parochialism and Publications Overload

Consumer policy is endemically multidisciplinary. You can make the argument as well as I. So I will practice parsimony and let you do it.

What appalls me--and I speak not only of the papers in this session but of many of the papers published in the Journal of Consumer Research, the Journal of Consumer Affairs, and the new Journal of Consumer Policy--is the parochialism that I note in the citation and revealed ignorance of adjacent disciplines. Consumer economists cite consumer economists, markets cite marketers, economists-economists, etc. It seems to me that this parochialism has great perils. First, we may omit essential elements and come up with inappropriate policy recommendations. An example: the consumer economists who advocated unit pricing and failed to recognize its information processing defects (Russo, 1977). Second, the limitations of narrowly conceived consumer policy research may become obvious to policymakers and hence our influence on policy may be greatly (and rightly) diminished.

It is appropriate to ask why this parochialism. One answer comes from the narrow, academic guilds into which we are organized and within which rewards are dispensed (Departments of Marketing, Consumer Economics, etc.). "Outsiders"--the adjacent disciplines--tend to be unrecognized, uncared for, and denigrated.

But a second answer comes from a concept that has

occupied students of consumer behavior much in recent years--information overload and, by extension, publication overload. This Conference, responding to the publish-or-perish imperative of the academic reward system, will contribute at least 600 pages more, certainly to be published and perhaps to be read. (Little wonder we tend to underread the literature of our adjacent disciplines!)

(I find it ironic and instructive that one of the proponents of the "Information overload" concept, Jacob Jacoby, did not discover and cite the articulator of an economic theory of information overload. I refer to Staffan Linder and his Harried Leisure Class (Columbia Press, 1970) in which he notes (Ch. 6) how increased affluence applied to fixed time will necessarily result in less search and increasingly sloppy decisions!)

We in consumer research have no monopoly on the publication overload problem; it is a scourge of all academia. The solution? Two steps, panaceas by no means, come to mind. The first is to make publication more costly, perhaps by charging a stiff review fee for submitted manuscripts. The second is to reward discriminating reading more, by counting, assessing, and by shifting the rewards in favor of discriminating reading, for example, by taking account in the making of salary and promotion decisions of the quantity and quality of review activities.

Information is so central to our profession that reform should start with us. More particularly, can we realistically expect policymakers to read the 600 or more pages that this Conference will produce?

#### Pilot Studies vs. Substantive Research

For empirical research novelty justifies the use of data taken from an intrinsically uninteresting, unrepresentative population, or a badly biased sample of an interesting population. In short, there is a place for pilot studies.

But substantive studies will deserve our attention only when they are based on adequate samples of interesting populations. As consumers of research, we must insist that researchers tell us enough, in the article we are reading, to enable us to make an informed judgment ourselves regarding the character of the population and the adequacy of the sample as executed. We should ask reviewers to dismiss as unacceptable (or only acceptable after revisions) papers that do not meet these standards.

#### A Matter of Proper Connections: The Conception, Execution and Follow-Up to Consumer Policy Research

Ask and answer an uninteresting or unimportant question and you have wasted resources. So it is a matter of some importance to pose a question that is worth answering.

Ask an important question and answer it unrealistically and you will also have wasted resources since your "answer" will prove impracticable.

Finally, answer an important question realistically and you may still fail, because the relevant decision-makers may fail to notice or to use your answer.

How may these pitfalls be avoided? One possible means is to touch base with consumer policy people of various roles before, during, and after the undertaking of research.

Whether any groups or individuals can take a sufficiently broad view of a field so as to be able to identify the

"proper" questions I am not sure. But many have tried. For those contemplating consumer policy research I would call your attention to several papers that seek to map the consumer policy field and its gaps.

A recent NSF Project (Consumer Affairs Institute, 1976) brought together an altogether remarkably assemblage of consumer advocates, consumer "bureaucrats," consumer affairs people from business, and consumer researchers to identify, through organized brainstorming sessions, important consumer research problems. A second NSF-sponsored Conference (Denney and Lund, 1978) included researchers, businessmen, consumer advocates, and academics and ex post sought their reactions to the consumer policy research that was reported at the Conference. A recent paper by a consumer economist (Shepard, 1978) has provided a series of taxonomies of consumer policy research that should help to identify gaps. Finally, responding to a call from the United Nations Centre on Transnational Corporations I, as one student of consumer policy, sought to organize the field and pose the most important issues (Maynes, 1978).

But one would be fatuous to expect that any group or any individual would raise either all the appropriate questions and certainly not the outlines of appropriate answers. Is it not precisely the original question or answer that is unlikely to be anticipated? By the same token, it can be argued that practitioners of consumer policy or consumer affairs are so close that they tend to be preoccupied with the immediate. Consider a possible policy to deal with the problem of consumer complaints: change the law so that complainants whose grievances are judged genuine are entitled to be compensated for the time they spend in seeking consumer redress. A plausible proposal, I suggest. But never have I heard it voiced by a consumerist!

As a final comment, I would like to reiterate my praise of Messrs. Ritchey and Claxton for incorporating the feedback of their research results to consumer policymakers. This is indeed a feature that all of us should seek to incorporate into our research programs!

#### References

- Arthur Best and Alan R. Andreasen, Talking Back to Business: Varied and Unvaried Consumer Complaints, Working Paper (Washington Center for the Study of Responsive Law, 1976).
- Consumer Affairs Institute, Toward the Development of a Comprehensive Consumer Research Program (Washington: National Science Foundation, 1976).
- Ralph L. Day (ed.), Consumer Satisfaction, Dissatisfaction and Complaining Behavior (Bloomington, Ind.: Indiana University, School of Business, 1977).
- William Michael Denney and Robert T. Lund, Research For Consumer Policy, Proceedings of a Conference (Cambridge: MIT Center for Policy Alternatives, 1978).
- H. Keith Hunt (ed.), Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction (Marketing Science Institute, 1977).
- E. Scott Maynes, "Consumer Protection: The Issues" (Unpublished Paper prepared for the United Nations Centre on Transnational Corporations, 1978).
- E. Scott Maynes, "Consumerism: Origins and Research Implications," in Eleanor B. Sheldon (ed.), Family Economic Behavior: Problems and Prospects (Philadelphia: Lippincott, 1973).

Daniel A. McGowan, Consumer Economics (Chicago: Rand McNally, 1978).

J. Edward Russo, "The Value of Unit Price Information," Journal of Marketing Research, 14 (May, 1977), 193-201.

Steve Salop, "The Noisy Monopolist,, Price Dispersion, and Price Discrimination," Review of Economic Studies, 1977, Forthcoming.

Sentry Insurance Company, Consumer at the Crossroads (Cambridge: Marketing Science Institute, 1977).

Lawrence Shepard, "Toward a Framework for Consumer Policy Analysis," Journal of Consumer Affairs, 12 (Summer, 1978, 1-11.

Hans B. Thorelli and Sarah V., Consumer Information Systems and Public Policy (Thorelli: Ballinger, 1977).

## "CONSUMER POLICY PROPOSALS" DISCUSSION PAPER

Hans B. Thorelli, Indiana University

The total range of consumer policy may be divided into consumer information, education and protection (Thorelli and Thorelli, 1977). In a general sense, consumer information comprises all data about individual markets and offerings. Consumer education is consumer civics. It provides the insights necessary to develop citizens into intelligent consumers. Consumer protection consists of measures taken by others than the individual consumer to safeguard consumer rights. It extends all the way from antitrust via complaints handling to product safety.

Consumer policymakers include consumer organizations, other citizen groups, business, government, educational institutions and the mass media. There are almost infinite tradeoff and reinforcement possibilities among information, education and protection, and among various makers of consumer policy in the struggle to enforce consumer rights and stimulate consumer responsibility. There is plenty of room for plenty of consumer policy proposals.

The paper by Goodwin, Mahajan and Bhatt describes the procedures devised by the Better Business Bureaus involved in arbitration, and cites data compiled from one of 100 such bureaus over a 4-year period. The paper by Moschis deals with the effectiveness of consumer-related courses in imparting consumer skills in grades 6-8 and 9-12, providing an empirical study of 806 students in 13 Wisconsin schools. The paper by Ritchie and Claxton presents the framework and methodologies in a large Canadian government-sponsored research project using consumers to identify shopping problems, to order the priorities among them, and to suggest solutions (with one service trade analyzed as an example), and how to go from sounds to things. In other words the Ritchie-Claxton paper constructs a consumer input scheme to generate consumer policy proposals.

### On Consumer Dissatisfaction: Consumer Arbitration as an Alternative Dispute Resolution Mechanism

It is a long, long way from consumer satisfaction to consumer disputes. Attempts to solve the problems in this field branch over all three parts of consumer policy. Broadly conceived consumer education by training in marketplace phenomena, decisionmaking, budgeting would better prepare consumers for the realities of product performance and dealer dealings. Consumer information if easily accessible and readily used would enable the consumer to choose among products and sources with discernment. Much attention is given to individual complaints handling procedures by policymakers in the legislative and executive circles in government, in businesses as well as in consumer-oriented groups, and of course in the academy, sometimes eagerly anxious to join in policymaking. Consumer protection has become the cry of the present.

The American Arbitration Association, since 1968 acting through its National Center for Dispute Settlement (NCDS), has played a key role in the development of all consumer arbitration programs. Particularly famed has been its special program with the Cleaners and Dyers Institute, now a quarter of a century in operation. New York City's Small Claims Court, which handles over 70,000 consumer complaints

a year, arbitrated more than 50,000 such complaints in 1974. Most cases are decided on the spot!

The innovative program by the NCDS in Albany is another extension of arbitration. Seven of Montgomery Ward's stores are precommitted to arbitrate any customer complaint that has gone through its grievance procedures and remained unsettled. A similar program exists in Harrisburg in cooperation with the Pennsylvania League for Consumer Protection and the Pennsylvania Retailers Association, with Sears, Bowmans and Pomeroy's participating. In the Cleveland-Akron office, the NCDS has arranged for the arbitration of consumer complaints referred by the Action Line, a local radio station, and the City Human Rights Commission. Some state attorneys general use arbitration procedures to settle disputes. The relatively new legal insurance devices appear to be edging toward arbitrating consumer disputes. Some local license boards have recently taken a more lively interest in complaints against businesses, such as home improvement itinerants. I have long advocated that such complaint-ridden service industries be required to post a performance bond and to submit disputes to arbitration.

So while arbitration in consumer disputes isn't such a new thing, it is always news when businesses join together to do something for the consumer. It would have been interesting to hear from our authors as to why the mechanism as developed by the BBB has not been used more. They might have come to this themselves had they cataloged the DISADVANTAGES of BBB arbitration as well as the advantages. One such reason they do give, however: the BBB itself has failed to educate the public and the businessman about the process. But there are others. There is the criticism as to whether the BBB itself should serve as the administrator of such programs as it is financed by business and has served as the mediator in the disputes going to arbitration. (The NCDS and AAA Regional Offices in Tucson, Cleveland, Pittsburgh and Miami conduct cases referred to it by the BBB.) Then, the program exists in only 100 places, primarily metromarkets. What happens to the nonmetro markets? Whereas there are plenty of disputes between consumers and the big companies, by far the largest number and the hardest to resolve satisfactorily nowadays are with the locals, the super offenders, where even a manufacturer has a hard time applying pressure. A Harris Poll apparently revealed that 44 percent of those who complain to the BBB remain dissatisfied, and 23 percent have a negative image of the Bureau as a result of the way their complaints were handled.

Another disadvantage is that there is a tendency on the part of some BBBs to hold back their offers of arbitration in cases deemed nonarbitrable. There is also the limitation of the program to marketplace disputes not including damages which might transcend the actual service or product complaint being arbitrated. Further several BBBs have set the limits on the dollar amount that may be arbitrated too low to be helpful to consumers. Do the BBBs know their 3 Rs: Refund, Repair, Replace? And, of course, overriding all these disadvantages is the reluctance of many businessmen to participate voluntarily in the process. As Table 2 reveals, over the 4-year period reported in one such bureau, in 35 percent of the instances where arbitration was suggested the business involved refused to take part. (Nationally, 4 out of 10 refuse.) While the time span from hearing to award is substantially shorter than a

full court trial (it would have been considerably more helpful to have the span from the initiation of the complaint), only 10 percent were completed within a month, while 40 percent took longer than three months.

This paper is a clear, succinct, well-written exposition of the BBB arbitration procedures and the experience of one bureau. There is, however, a great need to fit this arbitration procedure into a theory or a grand scheme of consumer redress. The tone is also a bit too enthusiastic; a more critical approach is essential in a research paper. Nationally, since the BBB arbitration program began 6 years ago, 2,300 cases have been heard. (No figures on how many resulted in awards.) Actually the BBBs are said to handle 1.5 million "complaints" a year, which would mean 9 million over the same six-year period. Suddenly 2,300 arbitration cases doesn't sound like much. Incidentally, it would have been very interesting to have been told the number of cases in the present study (was it embarrassingly small?).

Needed is a matrix for measuring arbitration as compared to other devices as a complaints handling mechanism. Such a set of criteria should be easily generated by these authors. As a beginning allow me:

Speed  
Credibility and fairness  
Local anchorage  
Consumer representation  
Product specialization  
Sanctions and their effectiveness  
Voluntary vs mandatory  
Number of complaints handled  
Number of settlements resulting in consumer satisfaction  
Number of settlements satisfying both buyer and seller  
Suitability as element of integrated complaint system  
Avoidance of business as well as consumer harassment  
Accessibility  
Inexpensiveness to consumers  
Consumer orientation  
Enforceability  
Extent of public awareness  
Total resource demands  
Informality  
Value of complaints handled

Many European countries--especially smaller ones like the Netherlands and Sweden--have developed quite effective procedures, using hearing boards composed of representatives from the industry involved and consumer representatives to decide on complaints. As many industry members belong to national organizations which collectively bind them to accept hearing board awards, enforcement of recalcitrants becomes a matter for civil law--violation of contract. In this country, it would be interesting to compare the experience and organization of BBB arbitration with that of MACAP, the major appliance industry consumer action panel.

#### Formal Consumer Education: An Empirical Assessment

There is a tendency to push to the schools anything that society has a hard time coping with (witness the discussions of sex education, marriage courses) and then complaining that the schools have failed. To evaluate the effectiveness of any educational program is difficult, despite the decades since Dewey and the countless doctoral dissertations of the schools of education collecting dust on library shelves.

Here the author has presented an empirical study designed to measure differences between those young people who have and those who have not taken many consumer-related courses in 13 schools in Wisconsin. He discovered, simply enough, that the differences were

not particularly great or significant. A serious deficiency is that most every subject had taken at least one broadly or vaguely consumer-related course.

In arriving at this conclusion the author learned a goodly number of interesting things. And like other important areas, consumer education may be bantered as well as battered around casually. He suggests, in addition to what is taught and how it is taught as possible and probable culprits that perhaps the students enrolling in such courses are those who make poor grades and gravitate to "pud" courses. (This at least used to be a matter easy to ascertain by examining the record or by asking the students themselves.) Of course, if this is true, consumer education is working. It brings the level of the "low" student to the level of the "high" student who did not take many courses. This is a fuzzy aspect of the paper. Could there have been a control group who took no courses? Were all courses elective? Were any obligatory?

Why not query whether the teachers are well selected for these courses. Health in my day used to be taught by the football coach. (He also taught social studies. It is a wonder he was not also into sex education.) In the area I live consumer education is an elective currently taught by the typing and shorthand teacher. Perhaps the author was not only at the mercy of school authorities in selecting his participants but also hesitated to query the teachers involved in the courses. What was the awareness and attitude of the teachers toward consumer education? Also what did the teachers say about the results of the study?

Missing is a scheme for defining, grading and weighting the wide range of the consumer-related courses, including (but not limited to) home economics, economics, environmental science, and career guidance. This may very well be the most serious problem with the study. Actually the author may have grappled with content analysis of some kind as he examined each "type" of consumer-related course and its relation to consumer skills (as well as age-group). However, this is not clear and neither is the author's definition of "type" of course. In perusing the tables one cannot but help noticing that even though some differences were statistically significant, they were nonetheless small.

The distance between a sixth grader (age 12) and a twelfth grader (age 18) is vast in terms of interest, concern, understanding--in short, in maturation. The tests to measure the skills appear too complicated for persons of 12 years, though Consumers Union claims even pre-school children should participate in consumer education. I am concerned that the consumer finance management question was not well thought through, especially as I am not at all certain that very many ACR members would come up with high scores.

On information-seeking, the method of counting sources and importance of information was in line with the current state of the art. It would have been exciting to see what greater role, if any, these information seekers play in disseminating information, using some sort of opinion-leader test (Thorelli, Becker, Engledow, 1975). As pointed out by the author, Table 2 has special relevance for social policy. If lower class students are to take advantage of the opportunity of mobility schools must energize them by particular efforts not least to change attitudes toward acquiring knowledge. This underprivileged group is likely not to move forward as long as their information is limited.

Consumer education is a hot topic. Publishers have been churning out materials haplessly. Companies such as J.C. Penney and Sears have serious, well-

constructed ready-to-go consumer kits, as does, of course, Consumers Union. So, too, does the HEW. Our neighbor to the north had devoted considerable resources to this matter. Scandinavia, especially Norway, has been successful in incorporating consumer education into the obligatory programs in high schools.

Lets give consumer education a real chance under clear auspices. As usual with empirical data of any consequence a multitude of questions remain unanswered, but not unanswerable!

#### In Search of Actions to Reduce Consumer Shopping Problems

Primarily methodological in nature this paper's strength lies in its systems approach. Using the results of a splendid extension of a management technique for program planning to the shopping problems of consumers combined with extensive interviews, the authors analyze data by conjoint and other measures and present one example in the automobile repair services. They close their paper by indicating procedures aimed at assuring the utilization or at least the awareness of the data by policymakers, principally in the government.

It is all too easy to think local in policy questions. And yet it is just in these questions that it is especially important to think ecologically. Problems may appear the same in all industrialized countries, but solutions rarely are--though superficially they may have much in common. Canada is a vast country, with 20 million inhabitants, a large immigrant population, plenty of nonmetro markets, and surely as many differences in outlook from the Maritime provinces to the Pacific northwest as exist in the U.S., even if the moon over Miami is the same as that over British Columbia. Maintaining a national policy in consumer matters has surely not been eased in recent years with all provinces, and not just Quebec, acting independently in a Canadian "province-rights" stance. This project is a noble attempt in adapting and extending methods to let the consumer speak for himself.

Figure 1 is of course far from complete. One important ingredient missing is the media. And then too it ignores that the one most important source of consumer information to the consumer is the manufacturer himself. Arrows in some cases need to go in both directions, such as lobby between government-suppliers and consumer information, consumers-government.

A truly severe omission is a full and clear presentation of the method of selection of the participants in the Nominal Groups; a complete description of the participants, other than upper/lower income, men/women; the number of groups and the number of participants in each group. In fact, the whole multi-phase project hangs on the sampling for the Nominal Groups. For a reader not familiar with NGT surely at least a paragraph should have been included. Some sort of justification for using five "major urban" centers, one in each of the five regions, should have found space here.

I consider refreshing the use of qualitative insights by average consumers, in the group context used here, regarding some critical problem dimensions in shopping. Too often professionals, including consumer professionals, look with disdain on lay inputs. For years I have recommended to industry the use of consumer panels--not to determine policy in the last analysis, but to put forward ideas and to help order priorities. Even many of the consumer organizations in Europe fail to see the need or value of consulting average

consumers.

Table 3 giving results of the analysis of the conjoint measures in the one aspect of the study reported is tantalizing. While space constraints explain why the data is so sparse, the lack of data nonetheless remains the real problem of the entire paper, and why as presented here it can only be viewed as a methodological piece and even incompletely documented as such. A fascinating prospect (apparently not observed by the authors) would be to use the results of this kind of study (as indicated by Table 4, for example) as the basis for broad-gaged public opinion polling. One of the results from our field studies of consumers in the U.S. and Germany--with excursions in Norway--was that average consumers far less than consumer advocates, but somewhat more than Information Seekers (the person who prepurchase-collects more data on products/services from more sources and values information more), look to the government and to regulation for solutions.

Of the five groups of product/services (only three of which were identified), the researchers picked for illustration the one where objective data is hard to come by, where local variations are the greatest and data collected may very well be valid for a short period only. This was unfortunate. Incidentally, studies from Sweden in the area of choice of automobile repair facilities indicate that the position of consumers has been greatly enhanced by the use of government inspection/diagnostic centers where they can get independent/objective information on the state of health of their automobile.

For some years Canada's Department of Consumer and Corporate Affairs maintained Box 99, a mailing address for consumers to contact for information or to lodge complaints. By 1972 20,000 complaints were registered annually and compiled in such a way as to allow for ready identification by problem, industry, locale. Even though the persons who availed themselves of this service were undoubtedly largely Information Seekers rather than your average consumer, to identify areas for attention this data bank could have been put to great uses. But as we say in the academy, never start with anyone else's data or for that matter his assumptions or method, if you can help it.

Nevertheless it is true to say that the authors have exhibited great ingenuity in transplanting and adapting for marketing research use a method of eliciting from consumers themselves how they view their shopping problems. It will be exciting to read the complete report.

#### Suggestions for Future Research

The National Science Foundation has funded several projects ostensibly devoted to consumer policy research proposals. One such project by the Consumer Affairs Institute listed 63 "research suggestions and considerations." Another, by the Center for Policy Alternatives at MIT listed 67 research priorities over three broad areas: research on consumers, research on policymakers and the policy process, and research on business practices. I suggest additionally a study on the impact (or lack of impact) of the research sponsored by NSF during its period of involvement with consumer research. Those heady days of lush grants may never return--someone would render a public service by examining what was accomplished.

Perhaps it is time to suggest that we should move beyond those research areas heavily oriented toward what I call consumer psychology--psyching out the consumer--and

the extensive use of sophisticated analytical tools on data of doubtful quality, sophomores of doubtful representativity, and topics of doubtful interest. We need to look to other disciplines, such as political science and organization theory (Thorelli, 1964). We also appear to forget that consumers constitute a key element of the market system as such. What are the effects of price controls on the consumer? How do laws/regulations affect consumers? What are the dysfunctional consequences of protective measures? There is an incredible underuse of court and Federal Trade Commission data. What are the consumer's problems in nonmetro markets? What are the consumer's strategies in the marketplace (when in doubt buy the most expensive, the least expensive; buy the branded item, the items that have been around the longest)? When does shopper tedium set in--and what conclusions may be drawn regarding channel design and city planning? What can standardization do for consumers (three sizes of toothpaste instead of twenty-six)? How much can metrication help the consumer in making choices?

It is easy to add to the kaleidoscope of research problems and policy issues. Surely we should be able to push comparative studies further along. What about the lot of the consumer under different economic systems? What is the cost to individual and society of the endless queues in Soviet stores and the endless bargaining in the bazaars of the Third World? A systematic comparison of the similarities and differences of the ghetto consumer and the consumer in developing countries has yet to be made.

Barely touched is the consumer education area. We need more study on clientele groups: aging, nonmetro, disadvantaged. As the paper here showed, some control of the content of consumer ed programs is probably necessary--and yet we know too little of what should go into them. In the area of product safety we need to face the tradeoffs between bans and regulations on one hand and informational programs and education on the other. In the area of consumer ethics, the need is great to get consumers to take marketplace responsibility after over a decade of emphasis on their rights. No dearth of research opportunity here.

There is a growing body of literature on labeling as a means of consumer information. Is it true, as often claimed, that labels are viewed by consumers as quality guarantees? Is there a difference in this regard between people who read the labels and those who merely see them? Would there be less misunderstanding in a combined labeling-certification program as advocated by us? (Thorelli and Thorelli, 1977). How effective is counter-information over time? Is consumer information overload a consideration in public policy? Researchers have discovered that some consumers actually avoid information. Who are these consumers? Do they make poorer marketplace decisions than others?

Of special value to business and consumer groups and other private consumer policymakers would be the development of standardized consumer satisfaction scales, analogous to the instruments of measuring employee satisfaction now routinely used by many progressive companies.

The nonmarket consumer participation in management decisions and in public and private consumer policy formation and administration is another area of interest. Corporate hot lines, consumer affairs departments and consumer panels are as much political institutions as they are market phenomena. To what extent may fruitful analogies be drawn between the labor movement and the consumer movement? What is the futurology of consumerism? To what extent can we expect consumer policy to develop on a pluralist basis

vs as a result of adversary proceedings before courts and legislative committees? How does one explain the phenomenon of Ralph Nader? Researchers interested in the study of the consumer movement or of consumer leaders in America should note the Center for the Study of the Consumer Movement, a documentary goldmine recently established by Consumers Union (but by no means limited to CU or CU leaders). If interested write Mrs. Sybil Shainwald, Director of the Center, c/o Consumers Union.

Speaking more broadly, it seems to me that academics can be helpful in the formation of private and public consumer policy in a variety of ways, such as:

- Problem definition
- Researching policy alternatives and attendant cost-benefit analysis
- Evaluation of policy performance
- Examination of dysfunctional effects
- Public policy forecasting
- Development of research and evaluation methodology

I am concerned that we don't do a good enough job of problem definition. To do this we need more systematic attention to the areas of futurology and environmental trends. A few pungent examples: what will be the consumer policy problems (and opportunities) emanating from such developments as

- A majority of women working outside the home
- Computerized consumer information banks
- Home computers
- Interactive TV
- The energy crisis
- Revival of downtown?

Clearly, in the crucial area of problem definition there is as much room for clear conceptual thinking as for empirical studies. This also applies to the development of coherent philosophies of consumer policy (here again comparative approaches would be helpful). Most countries have developed their consumer policy measures on an ad hoc, staccato basis. While it may be that this pragmatic-existentialist approach in the end is the only one feasible, an overall philosophy might be of considerable value in revealing inconsistencies as well as priorities and in pointing to possible dysfunctional consequences in other areas, such as individual freedom or the preservation of open markets.

We may have come a long way, baby, but there is still a long way to go!

#### References

Hans B. Thorelli, "Political Science and Marketing," in Reavis Cox, Wroe Alderson and S.J. Shapiro, eds., Theory in Marketing (Homewood, Illinois: Irwin, 1964).

Hans B. Thorelli, Helmut Becker and Jack Engledow, The Information Seekers: An International Study of Consumer Information and Advertising Image (Cambridge: Ballinger, 1975).

Hans B. Thorelli and Sarah V. Thorelli, Consumer Information Systems and Consumer Policy (Cambridge: Ballinger, 1977).



RESEARCH ON CONSUMER INFORMATION:  
PUBLIC SECTOR PERSPECTIVES

Steven E. Permut, Yale University

Abstract

The present study examines public sector perspectives toward research on consumer information (processing, dissemination, etc.). Three respondent groups were selected from seven federal agencies, state-level departments of consumer protection, and national consumer organizations. The "policy relevance" of consumer research, perceived consequences of information programs, and sources of research data were examined, along with constraints on research utilization in the policy making process.

Introduction

Consumer information programs and other efforts to require full disclosure of product/service information has increasingly attracted the attention of consumer researchers (cf. Ross, 1974; Day, 1976; Wilkie, 1975; Hutton, McNeill and Wilkie, 1978). These researchers bring with them a behavioral orientation to consumer analysis, coupled with a growing sophistication with statistical and computer-based research techniques and methodologies. One common denominator among researchers focusing on consumer information problems is their concern for a systematic, empirical approach to such questions as:

- What is consumer information? Can it be defined and objectively measured? (cf. Chaffee and McLeod, 1973; Hughes and Ray, 1974.)
- Does everyone benefit from increased information disclosure, or do only certain consumer segments benefit from such efforts? (Day and Brandt, 1974.)
- Does information disclosure help the consumer in making "better" or more "correct" buying decisions? Can one say that a consumption choice is really "better" or "more correct" than another choice? (Jacoby, Chestnut, and Silberman, 1976.)
- Is more information better? Or will too much information overwhelm the consumer, yielding dysfunctional consequences? (Jacoby, Speller and Kohn, 1974; Russo, 1974; Wilkie, 1974.)
- How should consumer information be displayed or otherwise communicated to the user? Do "information formats" vary in their ability to convey information more or less effectively? (Bettman, 1975; Russo, Krieser and Miyashita, 1975; Bettman and Kakkar, 1977.)
- Are there any long-term effects to be expected by information disclosure programs? Should such programs require parallel efforts in mounting consumer education programs? (Staelin, 1978; Bloom, 1976; Deutscher, 1973; Walker, Sauter and Ford, 1974.)

These and related questions continue to be raised by researchers addressing consumer information issues. Unfortunately, however, "answers" to these kind of complex questions are not easily resolved. While some progress has been made thus far, and some tentative findings have emerged, it seems unclear to what extent policymakers or consumer advocates will even consider

research findings--let alone use such research as part of their overall decision making activities (cf. Wilkie and Greyser, 1974; Wilkie and Gardner, 1974).

This study addresses the likely impact of research on consumer information as it relates to the decision making process of select public and non-profit sector spokesmen. It seeks to provide a comparative index of opinion and attitude toward research evidence as one element of the continuing thrust toward more complete disclosure of information in the consumer marketplace. It seems only reasonable to assume that research evidence that is not used or examined by those people responsible for influencing consumer information programs cannot provide important guidance from which informed perspectives can emerge. Specifically, the following were examined: opinions about consumer information disclosure, opinions about the contribution of research to policy making, perceived consequences of consumer information programs, and sources of information and research data available to policymakers in government and consumer affairs. Implications for more effective use of research dealing with consumer information issues are then discussed.

Method

Sample

Three separate respondent groups were identified as important for this study: first, directors or senior staff members of divisions and bureaus of federal agencies involved with consumer information disclosure programs; second, directors or senior staff members in individual state departments of consumer protection (often found within the Office of the State Attorney General); and third, senior spokesmen associated with national consumer organizations and public interest groups. Recourse to several published directories and source books defined an approximate universe of organizations and respondents in each of the three target groups.

Seven federal agencies were considered most active in the area of consumer information programs: Federal Trade Commission; Department of Health, Education, and Welfare; Department of Housing and Urban Development; Department of Transportation; Department of Agriculture; Federal Energy Administration (now the Department of Energy), and the Consumer Product Safety Commission. Over one hundred separate state departments of consumer protection when identified, most of which were associated with the State Attorney General's office. Thirty-five consumer organizations and public interest groups were identified as likely to be most active in consumer information efforts, although the extent of involvement, areas of expertise, or other relevant factors could not be determined in advance. The selection did include major national groups that would be likely to represent a reasonably wide sampling of those active in consumer representation efforts.

Questionnaire

Questionnaires were sent to 44 individuals specifically identified as directors or senior staff members within the seven federal agencies previously described. In

addition, 106 questionnaires were sent to the director or chairman of each state-level consumer protection agency (all states having more than one such agency). Finally, 28 questionnaires were sent to the director of selected national consumer and public interest groups. In all, 178 questionnaires were sent, each accompanied by an individually prepared cover letter requesting cooperation in the study. After two weeks, a reminder postcard was sent to all respondents with a third reminder sent during the following two weeks.

### Results

Usable questionnaires were received from 29 (66%) federal agency respondents, 43 (41%) state government respondents, and 17 (61%) national consumer organizations. Thus, an overall response of 89 (50%) was obtained.

How do respondents view consumer information disclosure issues? As shown in Table 1, federal, state, and consumer groups held fairly similar positions on a variety

of key issues. Highest agreement was expressed for the proposition that consumer information disclosure is in the public interest, regardless of whether the information is actually used by the consumer. The least amount of agreement was expressed for the position that more research is needed to understand how much information the consumer actually needs. Interestingly, consumer groups expressed the least amount of support for this type of research, while federal policymakers were more inclined to see a need for research in this area. Finally, it appears that concern for possible "information overload" as a consequence of too much disclosure in the marketplace is not viewed as a potential problem among the three respondent groups.

Table 2 illustrates the variation among the opinions held by federal, state, and consumer spokesmen toward the contribution of research to policymakers. In general, these respondents exhibit a clear scepticism, almost a distrust, of research on information disclosure topics.

In fact, about the best that can be said for research in this area is that it "raises new issues or perspectives"

TABLE 1  
OPINIONS ABOUT CONSUMER INFORMATION DISCLOSURE\*

Statements	Federal Government (N=29)		State Government (N=41)		Consumer Groups (N=17)	
	Agree	Disagree	Agree	Disagree	Agree	Disagree
1. Consumer information can be objectively defined.	79%	17%	80%	15%	88%	12%
2. Consumer information can be objectively measured.	72	21	71	24	65	12
3. Disclosure of consumer information tends to benefit all segments of the consuming public.	86	7	90	7	71	12
4. Required information disclosure helps the average consumer make "better" buying decisions.	86	3	93	5	94	6
5. In terms of the amount of disclosure, <u>more information is always better for the consumer.</u>	83	7	88	7	88	6
6. If too much information is disclosed, some consumers may become "overloaded" and not benefit as much from the information.	14	79	7	93	12	88
7. The manner in which consumer information is displayed (or communicated) can greatly affect its use by consumers	72	17	71	22	76	18
8. More research is needed to learn how much information is actually needed by the consumer.	76	17	63	24	59	24
9. Consumers differ in their abilities to effectively use disclosure information.	83	14	88	12	71	18
10. Consumer information disclosure requirements are in the public interest, whether or not the information is used by consumers.	90	10	90	10	94	6
11. The best one can expect from information disclosure programs is short term rather than long term effects on consumer buying behavior.	69	24	59	37	71	24
12. Consumer education efforts have been lacking in most previous information disclosure programs.	72	24	88	10	88	12
13. Consumer education efforts are a necessary part of effective information disclosure programs.	83	14	88	12	100	0
14. It is relatively simple to identify the type of information most consumers need to make informed buying decisions.	83	17	85	15	82	18

\*NOTE: Percentage of respondents not answering or uncertain have not been included in tabulation.

TABLE 2

## OPINIONS ABOUT THE CONTRIBUTION OF RESEARCH TO POLICYMAKERS

Statements	Federal Government (N=29)		State Government (N=41)		Consumer Groups (N=17)	
	Agree	Disagree	Agree	Disagree	Agree	Disagree
"Given your own experience and/or knowledge of research dealing with consumer information disclosure issues, would you say that..."						
1. the research is of high technical quality?"	45%	31%	54%	39%	76%	24%
2. the research is objective and unbiased?"	48	28	51	34	59	35
3. the research results can be generalized to the real world?"	41	38	56	37	53	47
4. the research has an acceptable level of reliability and validity?"	48	31	51	37	59	41
5. it is compatible with the positions and philosophy of policymakers like myself?"	34	48	46	41	53	41
6. it tends to be too academic and lacking in real-world relevance?"	48	38	39	56	35	65
7. it can be applied to the solution of consumer policy questions?"	45	41	49	41	53	47
8. it often contains explicit policy implications and/or recommendations?"	17	69	29	66	29	71
9. it is actionable (that is, can be realistically acted upon in some direct way)?"	21	66	34	59	35	65
10. it challenges existing ideas or assumptions about consumer behavior and/or consumer policy?"	34	41	39	46	53	47
11. it raises new issues or perspectives?"	62	24	63	48	71	29
12. it should be used by policymakers at all levels of government in dealing with information disclosure requirements?"	59	17	41	44	47	47
13. it should be used by policymakers in consumer organizations and public interest groups in dealing with information disclosure requirements?"	66	14	46	39	58	29
In addition:						
14. Most policymakers at Federal level are capable of judging what information the consumer needs to make informed buying decisions.	79	17	54	44	35	65
15. Most policymakers in state consumer protection agencies are capable of judging what information the consumer needs to make informed buying decisions.	66	24	78	17	35	59
16. Most policymakers in consumer organizations and public interest groups are capable of judging what information the consumer needs to make informed buying decisions.	59	38	71	24	82	18
17. Most policymakers in business are capable of judging what information the consumer needs to make informed buying decisions.	52	45	54	44	65	29
18. Most consumers themselves are capable of judging what information they need to make informed buying decisions.	62	34	68	27	71	29

for roughly two-thirds of the respondents surveyed. The strongest concern was apparently that respondents see research as not containing explicit policy implication and/or recommendations. All three groups are roughly split on whether research can be applied to the solution of consumer policy questions, perhaps due to the rather negative assessment of what respondents view as lacking in real-world relevance and "actionability."

In terms of the perceived ability of policymakers to judge consumer information needs, it is not surprising that each group views itself as most capable in this regard. In addition, the consumer is perceived as

slightly more capable than is business of properly judging the information needed to make an informed buying decision.

What are the expected benefits and consequences of information disclosure programs? Table 3 suggests that respondents from all three groups perceive positive benefits as a result of required information disclosure efforts--ranging from increased consumer sensitivity to nutrition, energy, and related issues, to improvement in the quality of products and competition. At the same time, respondents did not feel that increased product costs nor more limited product choices would result.

TABLE 3

## PERCEIVED CONSEQUENCES OF CONSUMER INFORMATION PROGRAMS

Statements	Federal Government (N=29)		State Government (N=41)		Consumer Groups (N=17)	
	Agree	Disagree	Agree	Disagree	Agree	Disagree
"Increased efforts to require full and complete information disclosure for consumers is likely to result in...						
1. an enhancement of the consumer's right to know."	97%	3%	100%	0%	94%	6%
2. improvement in the quality of products and competition."	83	17	88	7	71	29
3. increased opportunities for value comparisons."	93	7	93	7	94	6
4. increased purchase satisfaction."	90	10	88	7	94	6
5. increased consumer confidence in individual purchases."	90	10	88	7	94	6
6. an overall increase in consumer sensitivity to such issues as proper nutrition, energy conservation, product safety, and the like."	97	3	100	0	100	0
7. greater product standardization and therefore less product innovation and more limited consumer choice."	31	62	17	73	24	76
8. increased costs for products and services."	14	76	37	56	35	65
9. greater resistance on the part of business to voluntarily provide product information until <u>required</u> to do so."	66	34	24	76	29	53
10. the need for more enforcement activity on the part of regulatory agencies to assure compliance with disclosure requirements."	93	7	90	10	94	6
11. greater concern for consumer-based research to assess the effectiveness and benefits of information disclosure requirements."	52	48	27	73	29	71
12. greater need for educational efforts to equip consumer to effectively use information disclosure programs."	83	14	93	5	100	0

Consumer education was viewed as a necessary component of information disclosure programs, although the need for consumer-based research in assessing such efforts was less positive. Interestingly, nearly twice as many respondents at federal level concurred with the need for consumer-based research than was the case with state government or consumer group representatives.

And finally, respondents were asked to indicate their use of selected information sources, particularly those containing consumer-based research related to information disclosure issues. Although the sources listed in Table 4 are necessarily only suggestive of a larger number of similar resource documents available to the respondents, no previous study could be located that sought to index even a simple measure of comparative access and use for consumer research-oriented materials by policymakers. As Table 4 suggests, policymakers in each of the three respondent groups differ widely in terms of access and usage of potentially relevant research information. The Journal of Consumer Affairs, published by the American Council on Consumer Interests, received the highest overall access and usage rate (although federal agency respondents reported roughly half the level of use compared to state and consumer respondents). Perhaps most surprising is the relatively high rating of the Journal of Consumer Research, a relatively new journal published under the joint sponsorship of nearly a dozen professional organizations.

The level of access and usage of professional society proceedings, including attendance at their national

meetings, was highest for the American Council on Consumer Interests, followed by the Association for Consumer Research (ACR). Interestingly, 65% of the consumer representatives surveyed indicated at least one "use" of ACR proceedings and/or conferences during the previous 12 months, a finding all the more surprising given the level of technical sophistication normally found among ACR papers and conference meetings.

As one would expect, legal journals and other documents receive the major share of attention for most policymakers involved with consumer information programs. However, only occasionally do such publications provide consumer-based (or behaviorally-oriented) research information (although see, for example, Davis, 1977).

In terms of access to outside consultative and advisory sources, federal agencies claim the highest level of overall access. However, consumer groups appear to have a higher level of actual usage of professionals from departments of home economics, family economics, and similar areas, than found in state or federal government. Federal agencies report a moderate use of professionals from business and management departments. Whether these professionals were from economics, marketing, or other fields of management was not ascertained.

## Discussion

If one were to offer an overall profile of the perspectives of public sector policymakers toward research on consumer information, the following would emerge from

TABLE 4

## SOURCES OF INFORMATION AND RESEARCH DATA

Selected Information Sources	Federal Government (N=29)		State Government (N=41)		Consumer Groups (N=17)	
	Have access to	Used at least once	Have access to	Used at least once	Have access to	Used at least once
<b>JOURNALS</b>						
Journal of Marketing	24%	21%	7%	10%	24%	24%
Journal of Marketing Research	24	10	7	7	12	24
Journal of Consumer Research	48	34	39	37	59	71
Journal of Applied Psychology	17	7	5	5	6	12
Journal of Consumer Affairs	66	45	80	68	82	94
<b>PROCEEDINGS OF PROFESSIONAL SOCIETIES (or attendance at conferences/meetings)</b>						
American Marketing Association	14	14	5	12	18	18
American Psychological Association	7	17	2	5	6	6
Association for Consumer Research	31	31	15	24	53	65
American Council on Consumer Interests	48	45	46	49	88	88
<b>LEGAL JOURNALS AND SOURCES</b>						
Food, Drug, and Cosmetic Law Journal	97	69	98	76	88	71
Commerce Clearing House Reports	100	72	98	83	100	82
Prentice-Hall Consumer Credit Guide	93	24	83	29	76	24
Miscellaneous Law Journals and Law Review Articles	100	100	100	88	100	94
<b>Consultative &amp; Advisory Sources</b>						
College/University professors in:						
- departments of business and management	86	38	68	12	29	12
- departments of home economics	90	17	78	7	82	53
- departments of psychology	79	14	22	2	12	6

the present study: (1) respondents in the public sector expressed low support for more research dealing with consumer information needs; similarly, no real concern was expressed for the possibility of "information overload" occurring in the marketplace as information disclosure programs become more commonplace; (2) respondents exhibited scepticism toward the contribution of research on information disclosure issues, primarily, it appears, because such research is perceived as lacking explicit policy recommendations, as well as not being particularly "actionable"; and, (3) respondents in federal, state, and consumer categories view their own group (but not others) as being most capable of judging the real information needs of the consumer.

In addition to these findings, all respondents perceived positive benefits associated with information disclosure programs in general, as well as supporting the need for more consumer education efforts at the same time. Public sector respondents surveyed did, in fact, appear to have a moderate exposure to consumer-based research information, claiming access and usage of two well-known research journals along with their respective conference proceedings and/or attendance at one or more of their national meetings. This finding supports observations by Dyer and Shimp (1977), among others, whose research brings them into frequent contact with a variety of government policymakers.

One might reasonably ask at this point why the potential contribution of consumer-based research on information disclosure is not more actively embraced by policymakers. Clearly, the complexity of consumer behavior is not easily accounted for by any single discipline--such as law or economics, the two disciplines most frequently found in the educational backgrounds of our respondents. Yet the respondents surveyed here appeared relatively content to view consumer information policy issues from the primary vantage point of a lawyer or an economist. Why not from a behavioral perspective as well?

## Questions of Receptivity

While no single answer to the above question is likely to be satisfactory, it is useful to consider the possibility that lack of consumer research utilization may be due, in part, to the existence of a gap between consumer researchers and policymakers "...due to differences in values, language, reward systems and social and professional affiliations" (Caplan, Morrison and Stambaugh, 1975).

When respondents were asked for their views on the following statement, an average of 73% of all respondent groups "agreed" or "strongly agreed" with the proposition that: "A major factor affecting utilization of consumer-based research on information disclosure is a lack of mutual understanding and interaction between consumer researchers and policymakers." This high level of agreement is congruent with the findings of Caplan, et al., (1975), where 88% of their sample of 204 upper level policymakers in federal government "agreed" or "strongly agreed" with a similar statement related to social science knowledge utilization among policymakers. As the Caplan, et al., report concludes, "The need for reciprocal relations between knowledge producers and knowledge users in policy making positions is clear, but the problem of achieving effective interaction of this sort necessarily involves value and ideological dimensions as well as technical" (p. 29).

Given the parallel findings of the present study to the results reported by Caplan, et al., two additional questions were posed, each relating to the need for greater familiarity with the policy making process on the part of consumer researchers. Respondents were first asked to indicate their level of agreement with the following: "Consumer researchers won't be able to achieve effective knowledge utilization until they become thoroughly familiar with the policy making process." Across all respondents for the present study,

71% "agreed" or "strongly agreed" with this position. This is similar to the 73% agreement expressed by respondents in the Caplan, et al., study (again related to "social scientists" rather than "consumer researchers"). And finally, respondents were asked the following: "Consumer researchers tend to be naive about the political feasibility of the utilization of their findings." Sixty-nine percent of all respondents "agreed" or "strongly agreed" with this position; this compared with 78% of those respondents in the Caplan, et al., study.

#### Constraints on Research Utilization

Beyond the issues raised above, receptivity to consumer research on the part of policymakers must take explicit account of at least four interrelated factors that serve to constrain research utilization (cf. Weiss, 1976). These are:

1. Limitations of the research: including timing of the study (Hutton, McNeill and Wilkie, 1978); insufficient external validity (Wilkie, 1977), including issues of task environment, measures, and stimulus vehicles; inappropriate attention to demand characteristics (Sawyer, 1975); limited utility and generalizability to relevant population segments (Permut, et al., 1976); and, perhaps unavoidably, the problem of inconclusiveness of findings (either alone or in comparison to other studies).

2. Policy system obstacles: "system obstacles" faced by consumer researchers include unseen but real ongoing and reciprocal relationships with their counterparts in other agencies, legislators, professional staffs, client groups and others that compete for consideration in policy formation and recommendation. Further, policymakers may be looking for answers (preferably inexpensive solutions) that please the most people with the least disruption of the status quo. But perhaps the most important constraint herein is simply the universal assumption among researchers that policy decisions are "made" by some identifiable "decision maker." However, as Weiss (1976) points out, "Much policy doesn't seem to be made by a set of identifiable decision makers or by logical-rational procedures that could even take research into account" (p. 226). This argument becomes all the more important when viewed against the often-heard caveat that "...it is absolutely imperative that the researcher be in constant touch with policymakers to learn what their problems are and what types of research are most appropriate" (Wilkie and Gardner, 1974, p. 46).

3. Communication obstacles between researchers and policymakers: beyond the obvious problems created by "social science prose" for readers without social science research backgrounds, communication obstacles can be raised by inattention to translation of sophisticated models and analytic techniques into meaningful information for those trained in other disciplines. While technically-sound consumer research may follow narrative patterns of relevance to the research community, it will not be useful to non-researchers in this form. This includes, for example, greater attention to the order of presentation: government policymakers typically receive (and therefore expect) reports with recommendations and alternatives presented before the background and analytic discussions--the reverse of consumer research style.

4. Different "world views" of problem solving and decision making: including, under this broad umbrella, disparities in basic values and objectives, an "advocacy" tradition among policymakers, and differences in emphasis placed on "logic and rationality" versus "tradition and experience." Accordingly, for research that conflicts with intuitively held assumptions and

expectations, policymakers may well prefer to rely on other, more personal, data. In addition, some argue that research utilization in government decision making is often to support a predetermined position. Rather than mere "fact gatherers," policymakers inhabit a world of political reality that serves to mediate the potential impact of any research study, just as researchers inhabit another reality. Similar differences among private sector managers and management scientists have also been recognized in terms of research utilization, leading Hammond (1974) to argue that researchers must offer a form of "personalized rationality" that takes into consideration the needs and views of the decision maker toward the problem at hand.

#### Conclusion

On balance, it would appear that the effective use of consumer research by policymakers dealing with issues of consumer information can be enhanced on a number of dimensions. Foremost among these, it would seem, is an increase in interaction with, and a greater understanding of, policymakers and the policy making process. The expectation is clearly that greater interaction is the responsibility of the researcher, for as Wilkie and Gardner (1974) have noted, "public policy regarding consumer behavior is going to be made, with or without research evidence" (p. 46). Furthermore, in terms of consumer-based research on information issues, such as full disclosure of comparative product characteristics, labeling requirements, consumer education campaigns, and the like, greater efforts must be made by the research community to deal with and monitor the perceptions (or misperceptions) held by public policymakers concerning the potential relevance of research to the policy making process. Research on information processing of policy-makers themselves cannot continue to be overlooked if the desire to achieve policy relevance for consumer research is to be rapidly gained.

#### 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.
- Paul N. Bloom, "How Will Consumer Education Affect Consumer Behavior," in B. B. Anderson (ed.), Advances in Consumer Research, 3 (Cincinnati: Association for Consumer Research, 1976), 208-212.
- Nathan Caplan, Andrea Morrison, and Russell J. Stambaugh, The Use of Social Science Knowledge in Policy Decisions at the National Level (Institute for Social Research, University of Michigan, 1975).
- Steven H. Chaffee and Jack M. McLeod, "Consumer Decisions and Information Use," in Scott Ward and Thomas S. Robertson (eds.), Consumer Behavior: Theoretical Sources (Englewood Cliffs, NJ: Prentice-Hall, 1973), 385-415.
- Jeffrey Davis, "Protecting Consumers from Overprotection and Gobbledeegook: An Empirical Look at Simplification of Consumer Credit Contracts," University of Virginia Law Review, 63 (October 1977), 841-920.
- George S. Day, "Assessing the Effects of Information Disclosure Requirements," Journal of Marketing, 40 (April 1976), 45-52.

- George S. Day 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.
- Terry Deutscher, "Credit Legislation Two Years Out: Awareness Changes and Behavioral Effects of Differential Awareness," National Commission on Consumer Finance, Technical Report (1973).
- Robert F. Dyer and Terence A. Shimp, "Enhancing the Role of Marketing Research in Public Policy Decision Making," Journal of Marketing, 41 (January 1977), 63-67.
- John S. Hammond, "The Roles of the Manager and Management Scientist in Successful Implementation," Sloan Management Review, 15 (Winter 1974), 1-24.
- G. D. Hughes and M. L. Ray (eds.), Buyer/Consumer Information Processing (Chapel Hill: University of North Carolina Press, 1974).
- R. Bruce Hutton, Dennis L. McNeill, and William L. Wilkie, "Some Issues in Designing Consumer Information Studies in Public Policy," in H. Keith Hunt (ed.), Advances in Consumer Research, 5 (Ann Arbor: Association for Consumer Research, 1978), 131-137.
- Jacob Jacoby, Robert W. Chestnut, and William Silberman, "Consumer Use and Comprehension of Nutrition Information," Purdue Papers in Consumer Psychology, No. 163 (1976).
- 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 1974).
- Steven E. Permut, Allen J. Michel, and Monica Joseph, "The Researcher's Sample: A Review of the Choice of Respondents in Marketing Research," Journal of Marketing Research, 8 (August 1976), 278-283.
- Ivan Ross, "Applications of Consumer Information to Public Policy Decisions," in Jagdish N. Sheth and Peter L. Wright (eds.), Marketing Analysis for Societal Problems (Champaign, Illinois: University of Illinois Press, 1974).
- J. Edward Russo, "More Information Is Better: A Re-evaluation of Jacoby, Speller, and Kohn," Journal of Consumer Research, 1 (December 1974), 68-72.
- J. Edward Russo, Gene Krieser, and Sally Miyashita, "An Effective Display of Unit Price Information," Journal of Marketing, 39 (April 1975), 11-19.
- G. Sawyer, "Demand Artifacts in Laboratory Experiments in Consumer Research," Journal of Consumer Research, 1 (March 1975), 20-30.
- Richard Staelin, "The Effects of Consumer Education as Consumer Product Safety Behavior." Working Paper, Graduate School of Industrial Administration, Carnegie-Mellon University, February, 1978.
- Orville C. Walker, Richard F. Sauter, and Neil M. Ford, "The Potential Secondary Effects of Consumer Legislation: A Conceptual Framework," Journal of Consumer Affairs, 8 (February 1974), 144-156.
- Carol H. Weiss, "Using Research in the Policy Process: Potential and Constraints," Policy Studies Journal, 4 (Spring 1976), 224-228
- William L. Wilkie, "Analysis of Effects of Information Load," Journal of Marketing Research, 11 (November 1974), 462-466.
- William L. Wilkie, How Consumers Use Product Information: Assessment of Research in Relation to Public Policy Needs (Washington, D.C.: U.S. Government Printing Office, 1975).
- William L. Wilkie, "Consumer Information Processing and Public Policy Research." Working Paper Series, No. 2, Center for Consumer Research, University of Florida, June, 1977.
- William L. Wilkie and David M. Gardner, "The Role of Marketing Research in Public Policy," Journal of Marketing, 38 (January 1974), 38-47.
- William L. Wilkie and Stephen A. Greyser, "Consumer Research Inputs to Public Policy and Legal Decisions." Special Report, Marketing Science Institute, April, 1974.

CONSUMER RESEARCH INPUTS INTO PUBLIC POLICY  
DECISION MAKING: THE ROLE OF CANADA'S  
CONSUMER RESEARCH AND EVALUATION BRANCH

John L. Evans, Consumer and Corporate Affairs Canada  
C. Dennis Anderson, University of Manitoba  
L.G. McCabe, Consumer and Corporate Affairs Canada

### Abstract

In recent years, the Canadian government has taken positive steps to encourage policy-related consumer research. The ultimate objective is more effective consumer policy. This paper describes the organization and orientation of the Consumer Research and Evaluation Branch of Consumer and Corporate Affairs Canada. It also describes the policy development process employed by the Branch and Department and examines some of its ongoing activities. These activities are designed to improve interactions between policy officials and consumer researchers, and to facilitate the production of relevant and timely policy oriented consumer research. Finally, the paper looks into some future issues for policy research in the consumer field in Canada.

### Introduction

A number of authors have commented on the importance of consumer research inputs for public policy decision making (for example, Day, 1976; Wilkie and Gardner, 1974; Granbois and Olshavsky, 1972). The consensus appears to be that many policies are based on little or no consumer research. This situation can be attributed to many factors not the least of which is a lack of interaction between policy officials and consumer researchers.

The purpose of this paper is to describe some recent Canadian developments designed to facilitate the utilization of consumer research in policy decisions of the Canadian federal government. Particular emphasis is placed on the Consumer Research and Evaluation Branch of Consumer and Corporate Affairs Canada which is beginning to play an important role in focussing consumer research on policy needs.

### Background and Mandate

In order to place the following discussion in perspective, it is necessary to outline briefly the mandate of Consumer and Corporate Affairs Canada (CCAC) as well as the philosophy which has evolved for the implementation of that mandate.

Some people trace the initial impetus of a "consumer movement" to the early sixties and the statement of consumer rights, first articulated by U.S. President John F. Kennedy. In a period of increasing affluence and economic growth, the need was seen for measures to protect consumers against localized defects in the distribution system - the fly-by-night operator, hazardous products and misleading advertising - and to enhance the consumer's ability to choose effectively amongst alternative products on the basis of "accurate, honest and intelligible information". In Canada, these factors gave rise to the formation of the Department of Consumer and Corporate Affairs (now Consumer and Corporate Affairs Canada) in 1967. It is worthy to note that Canada was the first country to create a Department, headed by a Cabinet Minister, specifically charged with consumer protection.

The mandate of the Department is found in the Consumer and Corporate Affairs Act. As specified in sub-section 6(1), the duties of the Minister with respect to consumer affairs are to:

- "(a) initiate, recommend or undertake programs designed to promote the interest of the Canadian consumer;
- (b) co-ordinate programs of the Government of Canada that are designed to promote the interest of the Canadian consumer;
- (c) promote and encourage the institution of practices or conduct tending to the better protection of the Canadian consumer and co-operate with provincial governments or agencies thereof, or any bodies, organizations or persons, in any programs having similar objectives;
- (d) undertake, recommend or assist in programs to assist the Canadian consumer and be more fully informed about goods and services offered to the consumer; and
- (e) provide such inspection services for the protection of the Canadian consumer as
  - i) he considers necessary for the enforcement of any Act under his administration, or
  - ii) the Governor-in-Council may direct him to provide."

The consumer research mandate which is found in sub-section 6(2) authorizes the Minister to:

"undertake research into matters in which the powers, duties and functions of the Minister extend, co-operate with any or all provinces or with any department or agency of the Government of Canada or any organization or person undertaking such research and publish or cause to be published, or assist in the publication of, so much of the results of any such research as the Minister deems appropriate and in the public interest."

The provisions reflect the Canadian Parliament's intention that CCAC act with regard to problems of a broad nature as well as those which are specific to individual consumers. Generally, CCAC is charged with the responsibility for acting in the consumer interest irrespective of its form. However, while the departmental mandate provides for broad and far-reaching activities in the consumer interest, the early programs and policies of CCAC tended to take a much narrower focus due to the existence of a large number of rather specific, well-defined consumer problems. The approach to problem resolution in these cases tended to concentrate on the development of legislative and regulatory solutions rather than on the development of methods and procedures for improving the efficient operation of the marketplace through self-help and co-operative activities among consumer groups, governments and industry.

### Recent Developments

More recently, an awareness has developed within CCAC



and the Government that legislative and regulatory solutions to consumer problems may be less appropriate. First, the relatively straightforward problems have been dealt with. Due to their multi-faceted nature, many of the new and emerging consumer and social problems do not lend themselves readily to legislative or regulatory resolution. Second, changing economic realities have placed significant resource constraints on governments. These developments have created a situation which is no longer conducive to approaches to problem resolution which require extensive bureaucratic and regulatory structures. As such, CCAC has instituted a major shift in emphasis over the past few years. Specifically, policy now tends to place primary emphasis for problem resolution on the development of alternative actions designed to improve the functioning of the marketplace and on the consumer's ability to operate effectively within it. The result has been an explicit set of action priorities which emphasize: first, the enhancement of the consumer's ability to operate within the market system on his/her own behalf; second, the encouragement and promotion of effective problem resolution through consultation and co-operation with provincial governments, consumer groups, and industry; and finally, the development of programs for direct legislative and regulatory intervention in those areas where it can be shown that the identified problems cannot be resolved through either of the first two approaches, or that there is a basic market failure. This fundamental alteration in the operating philosophy of CCAC has had three effects. First, it has encouraged a much broader perspective with regard to the potential means for problem resolution, enhancing the Department's ability to achieve its legislative mandate. Second, it has placed increasing pressure upon program and organizational structures, originally designed to deal with the consumer interest in a narrower legislative and regulatory perspective, to discover effective ways of promoting informed decision making and self help by individual consumers. Third, it has led to a recognition of the need for a much more fundamental understanding of consumer behaviour in the marketplace prior to the development of government policy.

These effects have been reinforced by substantial social and political pressure directed at avoiding regulatory approaches to problem resolution and to seriously review existing regulatory structures with regard to their economic and social benefits and costs. There is an increasing realization that in many instances government regulation has created serious secondary problems. Indeed, these secondary problems have often been more severe, and, in many cases, more difficult to resolve than those for which such interventions were originally intended.

#### The Consumer Research and Evaluation Branch

In principle, the domain of consumer policy embraces the entire range of areas which are of interest to persons who are recipients and users of society's goods and services. The set of issues which touch upon the consumer interest is thus potentially unlimited. Ultimately, all final products of social and economic institutions are, in one form or another, items of consumption and all of the factors which effect this eventual utilization are of concern to the citizen qua consumer. Indeed, "consumer affairs" is not so much a label for a category of problems (although it has historically been that as well) as a particular perspective on traditional social and economic relationships: it looks at such relationships from the point of view of their final purposes, uses and consequences. Such a perspective makes it possible to define "the consumer interest" - often, but not always, by way of contrast to the "producer interest" - with respect to any particular area.

The very breadth of this perspective, however, has often frustrated attempts to develop an effective research effort in this area. There has been great pressure to develop quick solutions to specific problems; e.g., unsafe automobiles, adulterated and over-priced food, hidden credit charges, and the side effects of industrial production; rather than to develop a rational framework for analysis from which to approach and analyze emerging consumer issues. As such many policy responses have been ad hoc and have not accomplished their intended purposes. From a research, and indeed a policy, perspective what is most urgently required is just such a framework within which to organize available knowledge and determine investigative priorities. The development and implementation of this framework has assumed a high priority in the Consumer Research and Evaluation Branch.

In addition to the development of an appropriate research framework, it is necessary to ensure effective means for moving research results into appropriate policy and programme developments. In the early years of CCAC, the consumer research function was inward looking and oriented towards short-term problem analyses and program development. Very little attention was paid to systematic market intelligence, examination of long term trends, evaluation of fundamental or empirical research or to the value of intellectual capital building. Further, most research was performed in-house and little encouragement was given to the development of an effective and integrated consumer research community outside government. In 1975 this began to change. The purpose and functions of consumer research in the Department was reviewed and in late 1976 a new structure and philosophy was proposed and adopted.

#### Definition of Objectives

The first step was to carefully define, in an action format, the objectives of the research and evaluation functions within a consumer context. Essentially, the mandate of any such operation should be to 1) identify consumer problems, 2) investigate the nature and extent of these problems, 3) develop effective and efficient means for problem resolution, 4) develop plans for the implementation of policies and programs in conjunction with other interested parties, and finally, 5) evaluate the effectiveness of existing and proposed policies and programs.

Within this context, the objectives and means for problem resolution were stated as:

- (1) identification of current and potential problems likely to be of general concern to Canadian consumers as well as those likely to be of concern to particular classes of consumers such as the poor and the ill-informed;
  - by developing and maintaining an effective market surveillance system,
  - by maintaining close contact with governmental agencies within the provinces and outside Canada which are involved in consumer affairs, and
  - by attending and participating at conferences and meetings, both domestic and foreign, where consumer issues are likely to arise and where the consumer interest requires representation.
- (2) in-depth investigations and analyses of these problems, including reviews of legislative, regulatory and other actions taken in other jurisdictions;
  - by placing the problem in perspective with regard to its origin and fundamental causes,
  - by examining and evaluating the body of knowledge applicable to the particular problem,

- by the application of rigorous economic, statistical, sociological and behavioural analyses to the problem,
- by reviewing legislative, regulatory and program activities of other jurisdictions and evaluating the applicability and effectiveness of approaches taken in a Canadian context,
- by the preparation of thorough reports and papers on these problems and on the alternative methods for their resolution, and
- by circulating these papers and reports to obtain the benefits of the insights of other experts in the field.

(3) development of plans for remedial action suggesting policies which will effectively and efficiently resolve these problems and the presentation of these plans to senior departmental management;

- by ensuring that the objectives of any contemplated policies or programs are clearly stated in terms of their benefits to "end users",
- by undertaking socio-economic policy analyses which reflect sound research on the problem area,
- by preparing discussion papers and briefing materials which place the problem in perspective and suggest and evaluate alternative solutions,
- by establishing "action teams" to develop plans for the efficient implementation of plans and policies being considered, and
- by consulting widely with other federal and provincial government departments and agencies, consumer groups, industry, trade associations and other interested parties.

(4) implementation of these plans in cooperation with other branches, federal departments, provincial governments, consumer groups, industry, trade associations and other interested parties;

- by forming, where appropriate, Task Forces for policy implementation which include the various interested parties,
- by involving, to the extent possible, provincial governments, consumer groups and industry in the implementation and administration of programs, and
- by ensuring that sound operating procedures, consistent with the goals and objectives, are established for any policies or programs.

(5) evaluation of current and proposed legislation, regulation and programs in order to assess their effectiveness in achieving their stated objectives;

- by developing monitoring and performance measurement systems to ensure that policies and programs perform according to stated objectives in an efficient and effective manner,
- by conducting periodic policy and program evaluations to ensure their continued efficiency, effectiveness, and relevance, and
- by preparing policy reports and action plans for senior departmental management consideration.

These objectives and means clearly specified the role of the Consumer Research and Evaluation Branch in a functional form. They are stated according to their sequence in the policy development-implementation-evaluation process and include the main functions which are required for their realization. With this statement of objectives and means in hand, attention turned to the consideration of an efficient organizational structure and of an operational work plan necessary to achieve the reorientation of the Branch.

#### Phases in the Research Perspective

The combination of the breadth of perspective and the natural tendency of consumers and consumer groups to

focus on quick solutions to narrow and specific issues, coupled with an environment of economic restraint, tends to operate against planning and systematic policy development, and, therefore, against the establishment of a longer-term research perspective. This is not to say that a more systematic and comprehensive approach is impossible to achieve, only that its introduction must be phased and carefully planned. First, an overall plan for research must be developed which is broad enough to encompass the field, yet flexible enough to allow careful focus on individual issues. Second, a period of "capital building" must be undertaken to provide the perspective required in developing sound policies with regard to general as well as individual issues. In this capital building phase, there must be a serious review of priorities and decisions taken to postpone non-essential, short-term work in order to release required resources. Third, analyses of the "state of the art" in the various priority research areas must be completed and specific projects designed to fill serious gaps or to advance the level of understanding. Fourth, research personnel must be identified and retained who are capable of thorough analyses and who have a public policy perspective. Fifth, a series of position papers on the major policy areas must be prepared to act as the bases for future policy development and program design. These position papers must be reviewed and revised from time to time as new research results become available. During these phases, short-term priorities must be dealt with in as systematic and comprehensive a fashion as is possible. Indeed, the careful treatment of short-term problems can, and should, be integrated into the capital building exercise.

Implementation of the above process depends critically upon the specification of a functional structure within which consumer issues can be analyzed. To be effective, such a structure must be capable of encompassing issues along five main dimensions:

- (a) issues relating to subject-matter areas of direct involvement by consumers. These would include all issues arising from consumer products and services from the point of view of their design, production, pricing, purchasing, usage, safety and performance, as well as issues relating to the structure and operations of suppliers involved in the production and distribution process;
- (b) issues relating to the "modes" whereby distribution and trade are affected. These would include all issues arising within the realm of the exchange process in the marketplace, including packaging, physical distribution, and promotion;
- (c) issues relating to the selection and acquisition of goods and services by consumers. These would include all issues relating to the consumer decision-making process; pre-purchase search behaviour, information processing, product/service selection and post-purchase satisfaction, information and verification and conflict resolution;
- (d) issues relating to the legal and regulatory structure governing the activities of the various groups in the consumer system; and
- (e) issues relating to the impact of exogeneous or environmental factors on consumers. These would include inflation, energy and resource conservation, environmental protection, and so.

#### The Branch Organization Structure

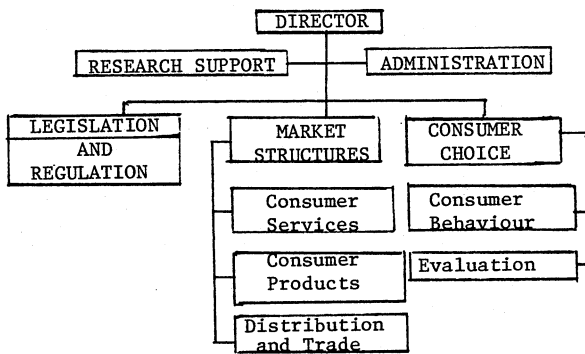
Within this context, a structure was developed (see figure A) which established three main focal points for

research activity: Market Structures, Consumer Choice, and Legislation and Regulation. The Market Structures Unit was sub-divided into three sub-areas: consumer products, consumer services, and distribution and trade. These three sub-areas were intended to encompass dimensions (a) and (b) above. The Consumer Choice Unit was sub-divided into two sub-areas: consumer behaviour and program and policy evaluation. These sub-areas were intended to encompass dimension (c)\*. The Legislation and Regulation Unit was undivided and was intended to encompass dimension (d). Due to the pervasive nature of the issues found in dimension (e), no specific research unit was identified. Rather, each of the three units identified above was charged with providing input from their individual perspectives. In addition to the specification of these three main focal points for research, an effective Research Support Unit was formed in order to optimize the productive use of time by research professionals.

This structure should not be construed to imply that each research unit operates in isolation. On the contrary, while the individual research units are charged with primary responsibility for issues arising within their particular dimension, it is required that for any given research project, a team made up of a member or members from all three research units be constituted to ensure that all aspects of a particular issue are considered.

FIGURE A

THE CONSUMER RESEARCH AND EVALUATION BRANCH  
CONSUMER AND CORPORATE AFFAIRS CANADA



The Operation Plan

Apart from the research function itself, a system for moving an identified problem through all the phases of the policy process was developed. In this regard, an action format was defined involving the entire Consumer Bureau of Consumer and Corporate Affairs Canada. It established clear means for identifying and classifying emerging consumer and market problems and for assigning responsibilities for action. Under this system, a Bureau Management Committee focuses on the problem recognition and problem assessment functions. Regular meetings of this management committee are convened to assess current and emerging developments that have been identified as adversely affecting the welfare of Canadian consumers. The problem assessment procedure is applied in an appropriately distinctive fashion to both new consumer market problems and those currently being addressed by Bureau policies and programs. The

\* It should be noted that program and policy evaluation functions are performed by each of the research units. However, since the majority of consumer programs and policies are designed to impact upon consumer decision

performance measurement, program review and policy evaluation systems, which are built into ongoing Bureau activities, provide a large part of the data utilized in the assessment of consumer problems under review.

Consumer problems are classified as falling into one of three general areas: (a) basic market failures; (b) substantive, recurring market imperfections and (c) exigent, small-scale market problems. Each of these initiates a different type of Bureau response. For example, if the identified problem is assessed as a basic market failure deriving from the inter-relationship of market structure variables, from the institutional and legal framework, from the performance of businesses or from consumer behaviour, and no current response mechanism exists, in-depth policy analysis is required in order to develop an appropriate government position. A policy analysis research plan is prepared, with periodic reports as the analysis progresses. The completion of the policy analysis research stage leads, in most cases, to the development and design of consumer programs and/or legislation. From this point on, the activities are similar to those for "substantive, recurring market imperfections", with steps involving a program implementation plan, a performance measurement system and a policy and program evaluation strategy. A matrix managed project team is assembled with research officers playing the lead roles at the policy analysis research stage, and program development and implementation officers taking up the workload during the program design and implementation phases. Periodic meetings of the entire team allow the project to enjoy the synergistic benefits of its multi-disciplinary composition.

The introduction of regular meetings of the Bureau Management Committee, with a carefully defined methodological approach to assessing the impact of Bureau activities on Canadian consumers, serves to center the time horizon of senior managers on the medium to long-term and facilitates the development of longer term program plans. The project team approach to the resolution of consumer market problems serves to rationalize the work of the Bureau and allows maximum exploitation of the comparative advantages of professionals in different organizational units.

The Conduct of Consumer Research

The above describes the structure and operations of the Consumer Research and Evaluation Branch from the organizational and process perspectives. An equally important perspective relates to the actual conduct of research, the Branch's role in stimulating consumer research outside Ottawa and its role in focusing the attention of researchers on important policy issues.

First the Branch has a policy of maintaining a relatively small staff (15) of highly qualified researchers trained in the fields of economics, sociology, behavioural sciences and law. These researchers have a heavy policy orientation in their work. Their role is: to identify and define emerging consumer issues and the nature of the research required by the policy process, to conduct certain research and to identify competent external researchers, to let and manage research contracts, to ensure that the research maintains a policy perspective, and to assist external researchers in translating their results for application by operating managers within the Department, or within other public policy units. The general approach to consumer policy is first to define and analyze consumer problems and develop

(cont'd.) making process in one way or another, the organizational responsibility for evaluations was placed with the Consumer Choice Unit.

appropriate policy responses, and second to determine which public policy unit can most efficiently implement that policy. In this process the respective governments maintain close contact through specially constituted Task Forces.

Second, the Consumer Research and Evaluation Branch has developed a number of means to facilitate consumer research in Canada. It has recently instituted a Consumer Research Newsletter which, while admittedly in its infancy, is designed to inform researchers and policy makers of each others needs and activities. Similarly, a Consumer Research Report publication series has been instituted to disseminate all major consumer research work done by or for the Branch. This is a refereed series with each publication being assessed by scholars in the relevant field. Of course, the series focuses attention on areas of current government interest. A Senior Management Seminar Series has been initiated which brings together researchers and departmental managers for the purpose of discussing the program and policy implications of in-house and contract research. These seminars are followed up by working contacts between managers and researchers to enhance the implementation of research findings. Conferences are being planned which will bring together government and business leaders, consumer representatives and researchers to expose problem areas and to discuss the options for their resolution in view of research results. These will occur infrequently when a major research program is in its final phases. At earlier times, contacts are maintained with business leaders and consumer representatives to obtain their inputs into research program design, as well as to elicit assistance in data collection, etc. A goal is to achieve a highly collaborative environment for problem resolution based on sound consumer research. Finally, contacts are being established with most major consumer research operations throughout the western world in order to share worthwhile results and methodologies.

Third, a research system has been developed to be used in defining major problem areas and implementing research programs. This system has been most successful in launching three recent research efforts. These are in the areas of consumer behaviour and decision making, energy conservation and products liability. Essentially, the exercise begins from the point of problem recognition and the priority assignment from senior management. With this assignment a small group of the principal Canadian researchers in the field are called together for a problem definition seminar lasting one or two days. Based on this a problem statement and a preliminary set of research program objectives are prepared. A full scale annotated bibliography in the relevant areas is then commissioned and the search for available data sources is begun. During this process the problem statement and research program objectives are revised in view of comments from outside experts and internal managers. Upon receipt of the bibliography, copies are sent to all interested researchers and preliminary proposals solicited for individual studies defined in the research program statement. A second research planning conference is then held to discuss proposals and finalize the research program. Final proposals are then developed and accepted. Typically the Branch attempts to plan a two or three year research program in some detail with quarterly updates.

This system appears to be providing an effective means for defining consumer issues, planning, conducting and coordinating consumer research in Canada, and for moving research results expeditiously through the policy process. Clearly there is much to learn and the system will undoubtedly be modified in many ways over the next few years. However, the initial results have been most promising and there is hope that developmental efforts will lead to enhanced consumer representation and pro-

tection in Canada, and to a greatly increased research input into the policy process.

#### The Way Ahead

While it is believed that progress has been achieved, there are a number of areas which should be mentioned as prime candidates for future activity. First, the newly implemented research-policy system must be thoroughly tested against organizational and political realities. Challenges to the success of the system will come on several fronts. For example, the ever present bureaucratic inertia must be overcome and the political need for quick, highly visible policy responses to complex problems must be dealt with. The actual operation must be carefully orchestrated to ensure a smooth and efficient transition of a problem through the policy development process. Careful attention must be paid to the coordination of the various organizational units which are involved in the systems' operation. Finally, due to the inter-disciplinary nature of the consumer field, it is critically important to harmonize the contributions of the disciplines of law, sociology, economics, consumer studies, marketing, applied psychology, and so on.

Second, a much improved problem recognition and definition capability must be developed. This will involve the integration of a large number of market surveillance functions such as the existing complaint and enquiry systems, periodic surveys of consumer satisfaction and market behaviour, and so on, in order to develop a policy oriented data matrix and reporting system for use by senior decision makers. More fundamentally, this capability will depend upon the development of a sound body of basic research in the areas of consumer behaviour, legal and regulatory structures, and the operations of consumer markets. In addition, it will be necessary to establish and maintain a network of researchers which spans the various disciplines identified earlier. These researchers must be provided with support, financial and otherwise, in order to stimulate interest in the field, concentrate their focus in areas of significant policy importance, and to ensure the efficient use of their available time.

Third, actions must be taken to improve the "total policy" concept. Attention must be given to the development of objective measures of the efficient and equitable operation of markets as viewed from the consumer's perspective. There must be a dramatic improvement in the scope and application of socio-economic impact analyses to policy alternatives. There is a need to enhance the role and importance of non-governmental parties in the policy-making process beyond the traditional "pressure group" system. Efficient alternatives to legislative and regulatory means for implementing policy must be developed. Approaches must be developed for concentrating the consumer's self-interest on problem resolution through information, education, the provision of effective redress mechanisms, and the promotion of advocacy both by the consumer and by third parties.

These are but a few of the topics which must be considered by consumer policy makers in the near future. We, in Canada, are hopeful that our efforts will combine with others throughout the world in the development of innovative and effective new policy responses to the resolution of current and emerging consumer problems.

REFERENCES

George S. Day (1976), "Assessing the Effects of Information Disclosure Requirements," Journal of Marketing, 40, 42-52.

D.H. Granbois and Richard W. Olshavsky (1972), "The Implications of Consumer Behaviour for Consumer Policy Decision," 47, 41-48.

William L. Wilkie and D.M. Gardiner (1974), "The Rate of Marketing Research in Public Policy Decision Making," Journal of Marketing, 38, 38-47.

CONSUMER PROBLEMS AND COMPLAINTS:  
A NATIONAL VIEW

Marc A. Grainer, Technical Assistance Research Programs (TARP)  
Kathleen A. McEvoy, King Research  
Donald W. King, King Research

Abstract

This study reports preliminary findings of a survey of a randomly selected cross section (N=2513) of the nation's households. Research issues addressed include: prevalence and generic types of consumer problems experienced, industries subject to consumer problems, degree of deprivation resulting from consumer problems, and complaining behavior. Results are compared with findings from other relevant studies that address these issues.

Since the early 1970's, the consumer satisfaction/dissatisfaction (CS/D) literature has expanded rapidly. Researchers have reported the results of numerous studies describing causes of (Anderson, 1973; Cohen and Goldberg, 1970; and Suprenant, 1977) and indices of (Day and Landon, 1976; Miller, 1976; and Pfaff, 1977) CS/D. Other investigators have studied levels of CS/D with selected groups of products and services (A.C. Nielsen Co., 1977; Best and Andreasen, 1976; Day and Bodur, 1977; and Handy and Pfaff, 1975).

During the same period, governmental decisionmakers and consumer advocates have become interested in this field of research. Their interest has centered principally on those studies where the findings deal with levels of CS/D by product/service category. Given the finite resources available for consumer affairs programs, effort must be channeled carefully into those problem areas of the greatest need. The level of CS/D with the various product/service groupings has been used in the setting of consumer protection priorities as a principal factor in determining the consumer population's "need for protection" (Day and Bodur, 1977).

In the past, the complaints received by consumer protection agencies have served as a surrogate measure for the general population's CS/D with products and services. There are, however, two significant problems with this practice. First, only a small percentage of such complaints ever reach governmental agencies or consumer advocacy organizations (Andreasen and Best, 1977; Best and Andreasen, 1976; and Day and Bodur, 1978). Complaints received by private manufacturers and retailers are generally not available for public review. Therefore, only a small, potentially unrepresentative sample of complaints is available for review. Second, the literature suggests that complaints do not constitute a representative sample of those problems experienced by the consumer population at large (Best and Andreasen, 1976; and Warland, et al., 1975).

Survey research techniques offer a much more direct and valid method for determining the level of CS/D with product/service groupings. The survey, with respondents selected on the basis of probability, allows the investigator to determine representative levels of satisfaction and complaint submission. While much survey work has been done in this area, most of the reported studies have concerned rather narrow lists of product/service groupings (A.C. Nielsen Co., 1977; and Handy and Pfaff, 1975). To date, CS/D with comprehensive listings of product/service categories has been measured in only two sets of studies, the Indiana University (IU) study (Ash, 1978; Day and Bodur, 1977; and Day and Bodur, 1978) and the Center for Study of Responsive Law (CSRL) study

(Andreasen and Best, 1977; and Best and Andreasen, 1976).

While each of these studies has made a significant contribution to the field, research design constraints limit, to some extent, the generalizability of their findings. The survey in the IU study utilized a probability sample of dwelling units in a single midwestern city. Therefore, as the researchers point out, generalization of their findings "to the national population would be inappropriate." In the CSRL study, the respondents were asked to evaluate their experiences with 26 products and eight services; a much less comprehensive listing of products and services than the nearly 200 items listed in the IU questionnaire. The CS/D findings are, therefore, limited to these 34 product/service groupings. Further, the 34 cities in which the CSRL telephone household interviews were conducted were not randomly selected. There was a substantial oversampling of households from six cities, and no rural households were included in the sample.<sup>1</sup>

This paper reports the results of a third survey that evaluates consumer experiences with goods and services generally. The sample for this survey consisted of a randomly selected cross-section of the nation's households. No category of product/service was excluded from the interview schedule.

The survey was designed and fielded as one component of a larger study of consumer complaint-handling practices conducted by Technical Assistance Research Programs, Inc. (TARP) under the sponsorship of the White House Office of Consumer Affairs (TARP, 1975; TARP, 1976a; TARP, 1976b; and TARP, 1978). The study's questionnaire was designed to provide the consumer affairs community with a data base for priority setting and policymaking. While the space available for this paper does not allow for a complete listing of the study's findings, the following questions are addressed:

What percentage of consumer households have problems with the goods and services they purchase?

Which generic types of consumer problems (false advertising, incorrect billing, etc.) are the most prevalent?

Which goods and services cause the most consumer problems?

What degree of deprivation is suffered as a result of consumer problems?

What percentage of the consumer households that experienced problems with the goods and services they purchased took some action to obtain redress?

---

<sup>1</sup>/CSRL presents a persuasive case that this oversampling of households does not bias the results. A high degree of stability is shown when responses by households from the six oversampled cities are compared with responses by households from the other 28 cities.

What percentage of the complaints submitted by problem households was "satisfactorily" resolved?

Parallel findings from the CSRL and IU studies are addressed where appropriate.<sup>2</sup>

#### Methodology

The data base for this study was collected via a nationwide survey consisting of personal interviews conducted with respondents found in a random sample of households. The survey sample was based on a national probability sample of households, consisting of 240 randomly chosen block clusters of households found throughout the continental United States. A sample of household addresses to be interviewed was chosen within each block cluster. A total of 4,327 households was selected by this design. Interviewers were given specific household locations as a starting point for each cluster, and they were instructed to follow a preselected route throughout the cluster. The survey was fielded in February 1976.

Of the 4,327 households chosen, 2,513 persons were interviewed, yielding a completion rate of about 60 percent. (There were 139 vacant homes and these are not included in the sample base.) Households not included consisted of 1,675 nonrespondents; 930 respondents not available for interviews after the initial visit and two callbacks; 714 potential respondents who refused to be interviewed; and 31 respondents who were not interviewed for other reasons. The interviewers questioned male or female heads of households, if available (or other adult household members, if not), about consumer problems they had experienced.

#### Level of Consumer Problems

A standard approach to measuring consumer dissatisfaction by survey is the "aided recall" method where respondents are asked how satisfied they were with listed categories of products and/or services purchased within a given time frame. The TARP study followed this approach with one significant modification. Instead of providing a comprehensive listing of product/service categories, respondents were handed a card containing a comprehensive list of consumer problems: ". . . here is a list of some of the problems (list provided in Table 1) that occasionally happen when people have bought or ordered things, received services, or when they try to get something repaired. Have you or anyone in your household had any problems like these in the past year? Which of these problems happened?" To make sure that any

---

<sup>2</sup>/When comparing TARP findings with CSRL and IU results, the differing time qualifiers used by these studies must be considered. In the TARP study, data were collected for all problems which occurred during the year preceding the interview, regardless of when the "defective" product or service had been purchased. In the CSRL and IU studies, dissatisfaction/problem data were collected for all products or services purchased within a defined timeline. If a consumer purchased an auto four years before the interview, dissatisfaction with a transmission failure that occurred six months before the interview would not be recorded in the CSRL and IU studies. This problem would have been reported by the TARP study. On the other hand, if the auto had been purchased two years before the interview but the transmission had failed 18 months before the interview, the problem would have been reported by IU but not CSRL or TARP. The CSRL study has a smaller dissatisfaction/problem universe than either the IU or TARP studies. The IU and TARP studies have overlapping but different dissatisfaction/problem universes.

problems not on the prepared list were counted, respondents were also asked: "What other problems not on this list happened?" Product/service data were then collected for each respondent-identified problem by asking: "What product or service was involved in this problem?" In sum, the order in which the product/service and type of consumer problem data elements were collected is reversed from the approach applied by the CSRL and IU studies.<sup>3</sup>

Before reporting the findings, it must be noted that this research only touches on the dissatisfaction end of the CS/D continuum. It does not follow that, because X percent of the households reported problems, the remainder of the sample was satisfied with the products and services they had purchased during the preceding year. The feelings of non-problem households probably ranged from mild dissatisfaction to very satisfied. From a consumer protection standpoint, it is the dissatisfaction/problem end of the continuum that is most relevant.

Of the national sample households, 32.4 percent (814/2513) reported consumer problems that had occurred during the previous year. These households reported a total of 1,582 problems, for an average of 1.9 occurrences for households reporting problems. The per-sample-household rate of non-price problems reported by the CSRL study of 2.4 (divided nearly equally between "strong" and "weak" problems) is more than three times greater than the 0.7 problems per-sample-household of the TARP study.<sup>4</sup>

---

<sup>3</sup>/This approach was followed for four reasons. First, the identification of generic sets of consumer problems is as important for the setting of consumer protection priorities as is identifying the industries in which the problems occur. This order of questioning allows for direct measurement of the problem dimension without the potentially biasing effect of a preceding industry screening item.

Second, there is a continuing debate in the CS/D literature concerning the inherent subjectivity and, therefore, ambiguity of responses to standard satisfaction survey questions (Hunt, 1977). (In the CSRL study, nearly as many non-price related consumer problems were reported by respondents satisfied with their past year's purchases as by those who were dissatisfied with the preceding year's purchases.) While the problem categories listed in Table 1 are certainly open to subjective interpretation, the ambiguity associated with this formulation is within more limited boundaries than the standard survey questions which are used to operationalize CS/D.

Third, the aim of the study was to identify what the respondents considered to be significant problems. While a "single incident recall" type question might unduly bias responses toward problems with the big ticket items, it was thought that an "aided recall" type question, which provided a comprehensive listing of product/service categories, would not provide enough discrimination between major and minor problems. The method used was a compromise thought to mitigate each of these potential problems.

Fourth, reasons of economy and interviewing technique foreclosed using the IU type of product/service listing. The questionnaire, which included many items not reported in this paper, would have taken too long to administer if the IU approach had been followed.

<sup>4</sup>/Non-price related problems are excluded from the problem base as being more a function of the general rapid escalation in inflation at the time of the CSRL study than of perceived industry malfeasance. "Weak" problems were identified by a probe: "How could it have

TABLE 1  
MOST PREVALENT TYPES OF CONSUMER PROBLEMS

Table 1 reports the percentages of problem households which experienced the various generic problems during the year preceding the survey. Three out of four of the most prevalent problems concerned basic deficiencies in the quality of products and services purchased. This is consistent with the findings reported by the IU (for services and intangible goods) and CSRL studies. Studies of consumer problems covering more limited product groupings (A.C. Nielson Co., 1977; and Handy and Pfaff, 1975) also support this trend.<sup>5</sup>

## Problem Industries

The eight industries in which products or services were subject to the most consumer problems were, in descending order of problem prevalence: automobile, appliance (radio and television), all other major and minor appliances, mail service, clothing, telephone, food and household items. In the automobile and appliance industries, less than half of the problems concerned marketing/business practices or basic product quality. Most of the problems reported were with the services provided by repair facilities. Both the CSRL and IU studies report a relatively high prevalence of dissatisfaction/problems in these two industries.<sup>6</sup>

Degree of Deprivation<sup>7</sup>

When setting consumer protection priorities, it is not enough to simply identify the most prevalent problem

been better for your household?"; asked of all satisfied purchasers. The "strong" problems, those identified without a follow-up question, are more comparable to the type of significant problems reported in the TARP study.

<sup>5</sup>The listing of generic problems as "aided recall" items elicited more detailed problem data in the marketing/business practice area (store did not have product advertised for sale, failure to receive delivery, etc.) than the open-ended follow-up probes utilized in the CSRL study. In the IU study, the "aided recall" probe for generic problems, which was administered as a follow-up question to respondents who were "highly dissatisfied" with a product or service, also elicited detailed marketing/business practice problem data. Such differentiation is useful for the setting of consumer protection policy. This data is of special relevance to agencies with enforcement and regulatory responsibilities.

<sup>6</sup>In the automobile and appliance industries, both the CSRL and IU studies report higher dissatisfaction/problem rates with repair services than with the quality of recently purchased items.

<sup>7</sup>Households reporting more than one consumer problem during the year preceding the survey were asked: "I'd like to ask you a few questions about only one of the problems that happened; that is, the most serious problem. Which problem was that?" When only one problem was reported, that problem was considered to be the household's most serious. The data presented in the following sections concern these most serious problems.

Type of Consumer Problem	Households Having This Type Problem	
	No.	%*
1. Store Did Not Have Product Advertised for Sale	203	24.9
2. Unsatisfactory Performance/Quality of Product (Workmanship/Ingredients)	182	22.4
3. Unsatisfactory Repair	165	20.3
4. Unsatisfactory Service (Unrelated to Repair)	127	15.6
5. Long Wait for Delivery	84	10.3
6. Failure to Receive Delivery	83	10.2
7. Overcharge or Excessive Price	78	9.6
8. Distasteful or Offensive Advertising	75	9.2
9. Product/Service Not as Ordered/Agreed on	72	8.8
10. Incorrect/Deceptive or Fraudulent Billing	70	8.6
11. Deceptive Advertising/Packaging/Pricing	66	8.1
12. Goods Received in Damaged Condition	65	8.0
13. Manufacturer/Dealer Didn't Live up to Guarantee/Warranty	58	7.1
14. Dealer/Salesman Misrepresented Product/Service	53	6.5
15. Failure to Receive Refund	43	5.3
16. Product Unsafe	30	3.7
17. Item Received Different Than One Bought	27	3.3
18. Product Harmful to Environment	26	3.2
19. No Return from Repair or Service	19	2.3
20. Instructions for Use/Care Unclear/Incomplete	19	2.3
21. Credit Terms Misrepresented	14	1.7
22. Unauthorized Repair or Service	13	1.6
23. Other	10	1.2

\* N = 814 households which experienced consumer problem(s); 1,582 incidences of problems



categories or the "offending" industries. In addition, some method must be utilized to weight consumer problems according to their relative "importance." (Given two problems of equal prevalence, a regulatory agency's efforts should be concentrated upon correcting the causes of the more "important" problem.) The CSRL study does this weighting by selecting the products and services to be included in its questionnaire. The IU study asks the respondents in its survey to rate the importance of each item they reported as purchased during the qualifying period. TARP attempted to formulate importance in terms of the actual or potential deprivation suffered as a result of the consumer problem. Three measures of deprivation were used: (1) actual or potential injury occurrence, (2) time lost from work, school, etc., and (3) potential financial loss.<sup>8</sup>

Of the problem households, 11.5 percent (94/814) reported that their most serious consumer problem had caused or had the potential for causing injury. Nearly 60 percent (55/94) of these households reported the potential for injury.

Of the problem households, 14.0 percent (114/814) reported lost time due to their most serious problem. Having to stay home from work, travel time connected with problem resolution, and waiting for repair/deliverymen were cited as the principal causes for lost time.

Respondents from each problem household were asked: "Would you or anyone in your household have been out of any money if the (most serious) problem was not corrected? How much?" Of the respondents, 61.9 percent (504/814) felt that a financial loss was possible. Of the 504 households that stated there was a potential financial loss, 451 specified the amount that could have been lost. The average potential loss for the 504 households that believed economic deprivation would occur in the absence of problem correction was estimated to be \$142. (Using the 814 problem households as a base, and, thereby, including the 310 households that did not foresee their most serious problem resulting in a financial loss, the estimated average loss would have been \$88.) This average cost estimate is the upward limit on loss that could be suffered as a result of the respondents' most serious problem. The range of responses was from \$1 to \$3,000.<sup>9</sup>

#### Action to Obtain Redress - Rate of Complaint Submission

With reference to the most serious problem reported, every problem household was asked: "Did you or someone in your household take some action to get this problem corrected or did you let it go and not do anything about it?" This question was worded to elicit reports of complaints (actions seeking problem correction). Sixty-nine percent (562/814) of the problem households

<sup>8</sup>/An index of relative problem deprivation, consisting of these and other measures of problem severity, could be developed for use in the setting of consumer protection priorities.

<sup>9</sup>/Estimating the upward limit of potential economic deprivation on the basis of responses to a single questionnaire item is imprecise business. An extensive series of follow-up questions would be necessary to clarify the ambiguities associated with this finding. (Did respondents include consequential damages in their calculation? Was the measure of loss based on original purchase price, replacement cost, cost of repair? etc.) The potential loss estimate must, therefore, be interpreted cautiously.

submitted (a) complaint(s) in an effort to resolve their most serious problem.

In the CSRL study, a less focused complaint question ("Did anyone in your household do anything about it [the 'strong' or 'weak' problem]?") was used. Unlike the TARP complaint question that was asked in reference to only the household's most serious consumer problem, the CSRL item was used as a follow up for all identified problems, serious or otherwise. CSRL respondents voiced complaints regarding approximately 40 percent of their problem purchases.<sup>10</sup>

However, when complaint rates were computed for "strong" non-price problems, the subgrouping of CSRL problems most comparable to the TARP study's most serious problem category, the complaint rate increased to 52.0 percent. The CSRL "strong" non-price problem complaint rates for appliance repair, automobile repair, televisions, and automobiles, the products and services reported by the TARP respondent households as being most often subject to problems, ranged from 75.0 percent to 67.6 percent. Therefore, the CSRL data are supportive of the TARP findings.<sup>11</sup>

In answer to the question: "Who was contacted first to try to correct the (most serious) problem?", more than 90 percent of problem households which took action reported submission of their complaint to the manufacturer or retailer. Less than 10 percent of the complaints were submitted to government agencies or private consumer advocacy organizations. Both the CSRL and IU studies report a similar reliance on the manufacturer and retailer to resolve consumer complaints.

#### "Satisfactory" Resolution of Complaints

Satisfactory resolution was defined from the consumer's point of view. The respondent households which initiated complaint action about their most serious consumer problem were asked: "Which of these statements (card containing statements listed in Table 2 shown to respondent) best describes your feeling about what happened as a result of your efforts to get the problem corrected?"

Table 2 indicates that more than 40 percent of the complaining households obtained a satisfactory resolution (more-than-asked-for, complete-satisfaction, and acceptable-solution assessment categories) to their most serious problem. More than 10 percent of the respondent households obtained "mixed" results (did-get-something assessment category) while more than 40 percent of the complaining households reported completely unsatisfactory results. Nearly 70 percent of the most serious consumer problems reported by the respondent households were, therefore, not satisfactorily resolved.<sup>12</sup>

<sup>10</sup>/Of the responses elicited by the CSRL question, 18.5 percent described such non-complaint actions as "exit" actions (e.g., modifications in buying behavior) and "self-help" (e.g., paying for a repair by independent expert).

<sup>11</sup>/In the IU study, 78.2 percent of the respondent households reporting dissatisfaction with service or intangible good purchases took some form of action. Of the actions, 53.6 percent of the actions would be classified as complaints according to the TARP definition. (According to the IU nomenclature, these were public actions; split evenly between requests for redress and complaints.) The IU study does not report a ratio between complaining households and problem purchases that is comparable to the complaint rates presented in the CSRL or TARP studies.

<sup>12</sup>/This rate consists of the sum of the not satisfied

TABLE 2  
ASSESSMENT OF EFFORT TO GET PROBLEM CORRECTED

Assessment	Percentage of Households Taking Action to Resolve Their Most Serious Problem*
I Received More Than I Asked For	1.1
I Was Completely Satisfied	21.4
I Was Not Completely Satisfied, but the Solution Was Acceptable	20.1
I Was Not Completely Satisfied, but I Did Get Something	12.8
I Was Not At All Satisfied	41.1
Other (Including Don't Know)	3.5
Total	100.0

\*N = 562 households which took action to resolve their most serious problem

In the CSRL study, 52.7 percent of the complaints about "strong" non-price problems were satisfactorily resolved.<sup>13</sup> The rate of satisfactorily resolved complaints for appliance repair, automobile repair, televisions, and automobiles ranged from 48.4 percent to 25.4 percent. The percentage of unremedied "strong" non-price problems reported in these four purchase categories ranged from 55.7 percent to 73.4 percent. These rates of satisfactorily resolved complaints and unremedied problems are consistent with the trends suggested by the TARP study.

#### Summary of Findings

The study's principle findings include:

32.0 percent of the national sample experienced one or more consumer problem during the year preceding the study.<sup>14</sup>

An average of 1.9 problems was experienced by these problem households.

Three of the four most prevalent consumer problems reported concerned basic deficiencies in the quality of products and services purchased.

The majority of consumer problems reported concerned marketing/business practices.

complainants (both "mixed" and completely unsatisfactory results) and the households which took no action to resolve their most serious problem divided by the total number of problem households.

<sup>13</sup>/Satisfaction was determined on the basis of responses to, "What was the result?", an open-ended question. On the basis of the reported method by which responses to this question were coded, the definitions of satisfactory resolution used in the CSRL and TARP studies are, in general terms, comparable.

<sup>14</sup>/Because the study was designed to investigate significant consumer problems, 32.0 percent problem household rate should be viewed as a very conservative estimate of the level of consumer problems generally. The rate most probably would have been higher if the questionnaire had been designed to also identify the more minor types of consumer problems.

The eight industries subject to the most consumer problems were (in descending order of problem prevalence) the automobile, appliance (radio and TV), all other major and minor appliances, mail service, clothing, telephone, food, and household item industries.

In the automobile and appliance industries, service-related problems predominated.

11.5 percent of the problem households reported that their most serious consumer problem had caused or had the potential for causing injury.

14.0 percent of the problem households reported lost time from work, school, etc. due to their most serious consumer problem.

61.9 percent of the problem households felt that a financial loss, averaging an estimated \$142, was possible if their most serious consumer problem was not corrected.

69.0 percent of the problem households submitted (a) complaint(s) in an effort to resolve their most serious consumer problem.

Most complaints regarding a household's most serious consumer problem were submitted to the manufacturer or retailer. Few were submitted to government agencies or consumer advocacy organizations.

More than 40 percent of problem households which initiated complaint action to resolve their most serious consumer problem reported totally unsatisfactory results.

Nearly 70 percent of the most serious consumer problems reported by the national sample were not satisfactorily remedied.

Because of the probability-based sampling design utilized for the study's survey, these findings have generalized applicability to the nation's households.

#### Discussion of Findings

When discussing consumer protection issues, the rhetoric of industry often consists of three defensive positions: (1) The number and seriousness of consumer problems suffered by the general population is not significant; (2) Only a small, vocal minority of consumers complain about the problems they experience with products and services; and (3) The great majority of those complaints about products and services, which are registered, are resolved to the satisfaction of the consumer.<sup>15</sup> The data do not support these contentions. What is suggested, instead, is the need for the private sector to upgrade its complaint-handling capabilities.<sup>16</sup> How such

<sup>15</sup>/It is often further argued that, even where industry is not totally at fault, a meaningful effort to resolve the problem will be made in order to retain the consumer's good will and future patronage.

<sup>16</sup>/Enlightened capitalism argues for business to develop techniques to identify those consumers whose problems have not generated complaints. While such consumers may not be unhappy enough to complain, they may be dissatisfied enough to change brands. Effective consumer problem resolution could prove a useful marketing technique.

upgrading may be most effectively accomplished is a worthwhile issue for future research.

## References

The consumer advocacy movement has long argued that there is a need for consumer protection. One policy adopted to protect such consumer interests has been the establishment of governmental agencies and private, voluntary organizations which, among other responsibilities, mediate disputes that consumers and business cannot resolve by themselves. The data document quite well the need for consumer protection. (Consumer problems exist in large numbers, and a sizable percentage of complaints is not satisfactorily resolved.) However, only a small percentage of households which complain about serious consumer problems seeks aid from the newly established third party mediation agents. Pinpointing the explanation for this latter finding is an important topic for future research. Explanations for this finding may include: consumer perception that these mediation agents are ineffective complaint handlers, lack of consumer awareness that these agencies exist, and/or the inability of these organizations to increase their caseloads without additional funding.<sup>17</sup>

The CSRL, IU, and TARP studies indicate the feasibility of utilizing survey data as a consideration in setting consumer protection priorities. The survey methodology provides the consumer affairs analyst with a data base generalizable to specified target populations and capable of pinpointing consumer problems by: (1) prevalence, (2) generic type, (3) "offending" industry, (4) importance (degree of deprivation), and (5) population segments at risk. To be put to its best advantage, the survey methodology should be utilized on a longitudinal basis. This methodology is appropriate for monitoring local as well as national consumer problem trends.

Critical review of the differing survey techniques reported in the literature is needed. Operational issues (order of items in questionnaire, telephone vs. personal administration of interview schedule, etc.) must be explored in depth before this social science methodology can be utilized to its fullest potential as an applied public policy technique.

---

<sup>17</sup>TARP's ongoing White House Office of Consumer Affairs survey of consumer complaint-handling practices in governmental agencies and private, voluntary organizations suggests that caseloads have become so heavy that additional funding would be required to significantly increase the volume of complaints presently being handled.

A.C. Nielsen Company (1977), A New Look at Consumer Attitudes Toward Product Quality.

Anderson, R. (1973), "Consumer Dissatisfaction: The Effect of Disconfirmed Expectancy on Perceived Product Performance," Journal of Marketing Research, 10, 38-44.

Andreasen, A.R., and Best, A. (1977), "Consumers Complain - Does Business Respond?" Harvard Business Review, 55, 93-101.

Ash, S.B. (1978) "A Comprehensive Study of Consumer Satisfaction With Durable Products," Advances in Consumer Research, Volume V, ed. H.K. Hunt, Chicago: Association for Consumer Research, 254-262.

Best, A., and Andreasen, A.R. (1976), Talking Back to Business: Voiced and Unvoiced Consumer Complaints, Washington, D.C.: Center for Study of Responsive Law.

Cohen, J.B., and Goldberg, M.E. (1970), "The Dissonance Model in Post - Decision Product Evaluation," Journal of Marketing Research, 7, 315-321.

Day, R.L., and Bodur, M. (1977), "A Comprehensive Study of Satisfaction with Consumer Services and Intangibles," Consumer Satisfaction, Dissatisfaction, and Complaining Behavior, ed. R.L. Day, Bloomington, Indiana: School of Business, Indiana University, 64-74.

Day, R.L., and Bodur, M. (1978), "Consumer Response to Dissatisfaction With Services and Intangibles," Advances in Consumers Research, Volume V, ed. H.K. Hunt, Chicago: Association of Consumer Research, 263-272.

Day, R.L., and Landon, E.L. (1976), "Collecting Comprehensive Consumer Complaint Data by Survey Research," Advances in Consumer Research, Volume III, ed. B.B. Anderson, Chicago: Association for Consumer Research, 263-268.

Handy, C.R., and Pfaff, M. (1975), Consumer Satisfaction with Food Products and Marketing Services, Agricultural Economic Report No. 281, Washington, D.C.: Economic Research Service, United State Department of Agriculture.

Hunt, H.K., ed. (1977), Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction, Cambridge, Massachusetts: Marketing Science Institute.

Miller, J. (1976), "Exploring Some Alternative Measure of Consumer Satisfaction," Marketing: 1776-1976 and Beyond, ed. K.L. Bernhardt, Chicago: American Marketing Association, 661-664.

Pfaff, M. (1977), "The Index of Consumer Satisfaction: Measurement Problems and Opportunities," The Conceptualization of Consumer Satisfaction and Dissatisfaction, ed. H.K. Hunt, Cambridge, Massachusetts: Marketing Science Institute, 36-71.

Suprenant, C. (1977), "Product Satisfaction as a Function of Expectations and Performance," Consumer Satisfaction, Dissatisfaction, and Complaining Behavior, ed. R.L. Day, Bloomington, Indiana: School of Business, Indiana University, 36-37.

TARP (1975), First Portion of Federal Central Office Phase of Contract No. HEW-OS-74-292, Feasibility Study to Improve Handling of Consumer Complaints: Evaluation and Inventory Reports, NTISPB-244182, Washington, D.C.: White House Office of Consumer Affairs.

TARP (1976a), Second Portion of Federal Central Office Phase of Contract No. HEW-OS-74-292, Feasibility Study to Improve Handling of Consumer Complaints: Evaluation and Inventory Reports, NTIS-PB-263081, Washington, D.C.: White House Office of Consumer Affairs.

TARP (1976b), A National Survey of the Complaint-Handling Procedures Used By Consumers, NTIS-PB-263082 Washington, D.C.: White House Office of Consumer Affairs.

TARP (1978), Consumer Problems, Consumer Complaints, and the Saliency of Consumers Advocacy Organizations: A Review of the Literature and the Results of a National Survey, Washington, D.C.: National Science Foundation (forthcoming).

Warland, R.H., Herrmann, R.O., and Willitts, J. (1975), "Dissatisfied Consumers: Who Gets Upset and Who Takes Action," The Journal of Consumer Affairs, 9, 148-163.

Stephen A. Greyser, Harvard Business School/  
Marketing Science Institute

All three of the papers in this session provide us with excellent approaches and useful information with respect to their subjects. While each deals with an independent area -- consumer problems and complaints, consumer research inputs into public policy decision making, and the role of consumer information programs -- all are linked by a concern for consumer input to public policy considerations. In the Grainer et al paper there is a comprehensive approach to gathering information on consumer problems and complaints as an underpinning to government (and business) action with respect to them. In the Evans et al paper there is a manifest concern for building consumer research inputs into public policy decision making. In the Permut paper there is a focus on how research on consumer information relates to the decision-making process in the public sector. All three of these papers address specific issues and problems of great pertinence. All do it in a stimulating way.

Let me comment on a number of specific points about each paper. In the course of these comments I shall offer both deserved kudos (of which there are a great number) as well as a few "slaps on the wrist" (which are relatively few in number). At the same time, I shall also raise some questions about the way the authors -- and we -- approach the issues and problems in this territory.

#### CONSUMER PROBLEMS AND COMPLAINTS

Grainer and his colleagues report on what is clearly a major study of consumer problems and complaints. Their treatment is comprehensive. The study is well-conceived and well-designed. The results -- even within the length limits here -- are carefully reported, particularly the comparisons with other studies. Further, there is strong sensitivity to measurement problems reflected in the research and the report.

Both in this report and in much of the other research on consumer satisfaction/dissatisfaction, we find a common concern for identifying those problem areas that are of greatest importance. A variety of definitions of "greatest importance" exist. The TARP research should be commended for incorporating both an incidence measure and a measure of degree of deprivation (cost/risk) as indicators of importance. All of us should be aware that there are other measures of salience that are pertinent in assessing consumer problems -- e.g., those problems which people remember most clearly, those which generate a high degree of annoyance, etc. My point here is simply that multiple definitions should be incorporated in research of this type -- perhaps even more of them than were included here. An additional point about measurement concerns the action taken by the consumer. Beyond the specific inquiry with respect to taking action, I would like to see a question for those who did not pursue the matter as to "why not?" The responses to this question may tell us something more about salience. They may also tell us more about the ways in which current complaint-handling and complaint-resolving procedures are adequate or not. When a consumer says "It isn't worth the effort," he or she may be telling us that for certain kinds of product categories/kinds of complaints/dollar values involved it is not worth it in terms of the consumer's own view

of the value of his or her own time or money. But we may also learn about where the weaknesses are in current systems with respect to "ease of entry" of one's complaint and/or likely resolution thereof.

Two important points raised in the concluding sections of this paper warrant reinforcement. The first of these is the importance of employing a survey methodology on a longitudinal basis in order for it to be put to its best advantage. Second, the authors are correct when they point to the usefulness of exploring operational issues in methodology in this area at this point -- e.g., item sequence, mode of administering interview, etc. In my opinion concern for these kinds of issues is another indicator that the consumer satisfaction/dissatisfaction field has reached a point at least of healthy adolescence, if not of full maturity. Again, I believe that multiple approaches are appropriate -- in order to get a better sense of calibration of the instruments that are being employed. Further, as the amount of baseline data grows, the utility of the data for policy making grows along with it.

A number of questions -- and some observations -- arise whenever I read or listen to a well-done piece of research in this area. (It is only the well-done pieces of research that permit one the time to raise these more basic questions!) Much research of this type concentrates on a perspective from the dissatisfaction or problem end of the spectrum only. Certainly -- at least from the public policy viewpoint -- this is an appropriate perspective. At the same time, those doing research from such a perspective should be aware of the fact that one of the elements that puts certain kinds of problems and dissatisfactions into an even broader perspective is the degree to which there are satisfactions and nonproblems with the particular product category, service, etc. In my view at least, business concerns with respect to satisfaction/dissatisfaction -- and business certainly is very interested in this field even absent the public policy concerns -- should take into account the broader perspective just mentioned. In a sense, for certain kinds of problems, we might expect a "normal level" perception of problems or dissatisfaction to exist. In other words, no matter what efforts (by business and/or government) are made at improving complaint handling or complaint resolution there will always be some proportion of people who will respond to the TARP-type questioning about degree of satisfaction with the ultimate resolution of their problem as "I was not at all satisfied." In essence, I am simply citing the phenomenon of elevated expectations that often accompany improvements in substance or process in any area. For example, in our research some years ago on historical attitudes of Americans toward advertising<sup>1</sup>, it was intriguing to discover that data seemed to run roughly in parallel on public attitudes toward truth in advertising and toward the standards of advertising: more specifically, even though questions over the years indicated that a large proportion of the public believed that advertising standards were improving, at the same time questions with respect to the truth quotient in advertising showed a roughly consis-

<sup>1</sup>Stephen A. Greyser and Raymond A. Bauer, "Americans and Advertising: Thirty Years of Public Opinion," Public Opinion Quarterly, Spring, 1966.

tent modest plurality saying that advertising did not present a true picture of the products advertised. Cognizant as we were that truth and standards are not precisely congruent, nonetheless the only conclusion we could draw from the consistency of the data over many years was that the public's standards in fact were rising. I believe that the same phenomenon has been occurring and will continue to occur with regard to consumer problems and complaints.

A final point. Personally, I continue to be amazed by what to me are relatively low percentages of households reporting problems of the sort explored here when one considers the vast number of products and services purchased in the course of a year by the typical family in this country. In the TARP study, about one-third of the households reported consumer problems having occurred in the past year. In the Sentry study of American attitudes toward consumerism, we learned that this percentage can be increased by asking simply whether "In the past year or so has there been a time when you wanted to complain about a product or service you have purchased or hasn't that happened?" In response to this question, 63% of a national cross section said that there had been a time. Only 47%, however, answered "yes" to a related question as to whether they "actually did complain." When looked at in conjunction with the TARP data and with Nielsen data -- and when viewed in the broad context of literally millions of transactions taking place annually -- I am surprised that the percentage of households who would report a problem -- maybe even a salient problem -- is not very near to 100%! By noting my reactions to these data, I do not mean in any fashion to trivialize the importance of the kinds of problems that consumers indeed do have. From the vantage point of both government and business, we must know about the incidence and the seriousness of the problems that the public says it experiences. Research such as that reported here goes far to offer us a well-honed approach to discerning what these problems are. In turn, that helps both government and business make more informed judgments as to which kinds of problems and issues call for enhanced attention and improved resolution -- and by whom.

#### PUBLIC SECTOR PERSPECTIVES ON CONSUMER INFORMATION

In the wake of a growing amount of consumer behavior research ostensibly addressed to the concerns of public policy makers or consumer advocates, Permut appropriately explores the question of the extent to which these groups "will even consider research findings -- let alone use such research as part of their overall decision-making activities. In a real sense, he raises a number of important questions about the desirability and utility of research information (such as that being presented at much of this conference) from the vantage point of those in the public sector. He appropriately turns his attention to the perceived information needs and the information processing of policy makers themselves. As he notes in his concluding paragraph, he is fully aware of the cogent and pragmatic statement of Wilkie and Gardner that "public policy regarding consumer behavior is going to be made, with or without research evidence."

Despite the importance and the intriguing character of the investigation, there are some key questions that must be raised about the research itself. These are not surprising to any of us engaged in attitude research. One concerns the 50% response rate, particularly with such small samples. A second concerns the small cell sizes in the three individual response groups. A third relates to the validity of gathering information of this sort via an attitude questionnaire without personal interviews. A fourth is the serious question as to whether "consumer groups" are part of the "public sector" in the same way that federal government and state gov-

ernment people are. Rather than batter the research on each or all of these points, I would prefer simply to call your attention to these important questions. I suspect that each of you in the audience can easily contribute at least one paragraph of amplified critique with respect to each of these points!

Thus I would like to turn my attention to the nature of the findings -- however tentative they may be -- and what I think they suggest with respect to those in the government sector and in consumer groups. Before I do so, let me note that the specific subject area selected for examination -- namely, consumer information programs -- is a particularly pertinent one. Public policy proposals with respect to mandated information disclosure are rife. Researchers in the consumer behavior field are conversant with the contentions around the question of "information overload." Thus, opinions from those in the public policy sector about consumer information disclosure and the perceived consequences of consumer information programs are of interest.

Turning to observations on some of the findings, I think it is very revealing of the value system of those in the public sector that there is very strong disagreement on the part of all three respondent groups regarding the possibility of information overload. Virtual unanimity is shown by each responding group with respect to the propositions that required information disclosure helps the average consumer make better buying decisions, that more information is always better for the consumer, that consumer information disclosure requirements are in the public interest whether or not the information is used by consumers. As someone who was involved very early in discussing the value systems and "views of the marketplace" held by those in the public sector (and by those in business), I find the strong consensus on these issues to be fascinating -- even if not surprising.<sup>2</sup> These are the findings in Table 1 of the paper which I think are most interesting. (Let me note a point of interpretive disagreement with respect to the data on whether "more research is needed to learn how much information is actually needed by the consumer." While it may be true as Permut states that "the least amount of agreement was expressed" for this proposition, the data themselves show a strong positive endorsement for the proposition.)

The next section of the paper focuses on concerns with respect to relevance and usability of research. Here again, I find the data very revealing of the value systems of those involved. In the table below I have reconstructed data from the study in terms of how each responding group views the capability of different people or organizations to judge the information consumers need to make informed buying decisions. In my chart, I have taken each responding group's view of its own capabilities as part of a single row entitled "my organization's capability." What is seen here is that in every case the responding group considers its own capability to be much better than that of "most consumers -- and in every instance each responding group considers its capability to be even greater than that of business. Perhaps I am easily amazed ... but I find this to be more than a simple statement of reinforcement for one's own position and the validity of one's institution. Rather, I find it a manifestation of arrogance -- reflective of the worst dimensions of "big brother" government, and technocratic expertise. [At this point I shall pause and let you reflect on whether I am easily amazed, whether I should simply have said, "this is what one should naturally expect" and simply shrugged

<sup>2</sup> Raymond A. Bauer and Stephen A. Greyser, "The Dialogue That Never Happens," Harvard Business Review, November-December, 1967.

my shoulders, or whether I have an unsupportably lofty view of consumers' own capabilities and business' capabilities to learn about consumer information needs.]

TABLE 1  
VIEWS OF GROUP CAPABILITIES

	Respondent Group		
	Federal Government	State Government	Consumer Groups
My organization's capability	79%	78%	82%
Most consumers' capability	62	68	71
Business' capability	52	54	65

Still looking at Table 2 of the paper, I am intrigued that consumer groups believe that neither federal nor state government people are as capable of judging consumer information needs as are consumers and business -- let alone consumer groups themselves, which are (naturally) at the highest level of insight. This suggests to me -- to return to an earlier point briefly -- that consumer groups should not be considered analogous to the government groups as part of the "public policy system" with regard to this issue (and other consumer issues as well).

Since Permut's data show that people in the public sector think that information disclosure is an unmitigated good (even to the point of denying increased costs of same), it is interesting to examine their own information sources with respect to consumer research information. In this context again we must note a problem with the sample -- in this case a problem in sample selection, rather than in response rate. The people questioned are those who are "directors or senior staff members" in government and "senior spokesmen" in the consumer groups. These people -- compared with more junior staff people -- are in my opinion least likely to be attentive to what is going on in the consumer behavior research field -- for the usual reasons of the character of their responsibilities, the nature of their early training, etc. Since virtually none of the respondents has any formal background in consumer behavior and consumer behavior research, it is not surprising that Permut finds that they are "content to view consumer information policy issues from the primary vantage point of a lawyer or an economist." The inference I draw from Permut's data on "sources of information and research data" (Table 4) is that consumer behavior people should try to penetrate the law journals more extensively than we have so far!

I find Permut's conclusions with regard to utilization sensible with respect to both policy makers and researchers. All of us need to have reinforced the facts about the world in which public policy makers live. This is particularly true with respect to the style of that world -- dimensions such as political vs. substantive considerations as having leverage in the system, the time span within which policy must be made and actions taken, the "translatability" of research as an important consideration in developing policy and action. Much of the resistance to using research for policy formation (in contrast to looking for supportive research data) grows from an adversary approach for one's job. As Permut notes in his fourth constraint on research utilization, the advocacy tradition in the public sector is a basic value that is largely incompatible with research-based policy formation. Indeed the "I know what is best for the public" philosophy that seems to pervade the responses represents an even larger impediment -- even if one were to overcome the

advocacy tradition!

As I have noted already, the major contribution that this paper makes is simply its subject area: research on information processing of policy makers themselves. Certainly that examinations of both the values and the policy making behaviors of these groups -- especially using more clinical rather than survey techniques -- will become more frequent among researchers with an interest in this field. I cannot resist a closing comment with only a modest quotient of facetiousness. As researchers, we raise the question as to whether more consumer information will lead to more or less actual information use by consumers. Perhaps we should add to the questions in this area whether more consumer research information available to the public policy sector will lead to greater or lesser use of that research by people in that community!

#### CONSUMER RESEARCH INPUTS INTO PUBLIC POLICY DECISION MAKING

I've chosen to comment last on this paper because it is "where it all comes together." This is a very important paper simply because it treats the rationale and objectives of trying to build consumer research inputs into public policy decision making in a formal organizational and operational manner. It is in this area that I find the paper to be most worthwhile.

First, I would like to reiterate five points in different parts of the paper. (In so doing I am conscious that the authors (and I) are "preaching to the converted.") These five points are:

- Policy-related consumer research that is relevant and timely is highly desirable.
- Understanding consumer behavior in the marketplace prior to [my emphasis] developing government policy is desirable.
- Life in the public policy sector generates great pressure to develop quick solutions to specific consumer-related problems rather than research-based policy formation.
- There is a need to move research results into policy and program development.
- There is a need for research "capital building" in order to provide perspectives for sound policies.

At this point again I leave it to each of us in the audience to nod vigorously and amplify any or all of these points with a silent paragraph or two!

Overall, this paper presents a sensible rationale for the new (1976 ff.) role of the Consumer Research and Evaluation Branch. It presents logical and reasonable objectives for research. It indicates patterns and processes which seem sensible to me -- at least on paper. Its classification of three broad typologies of consumer problems (basic market failures, substantive recurring market imperfections, and exigent small-scale market problems) strikes me as appropriate to the different kinds of Bureau responses indicated.

Beyond this sensible description of the structure and operations of the Branch from the organizational and process points of view, there are additional useful elements in the CREB system. These include the ways in which the Branch tries to focus the attention of outside researchers on important policy issues -- as well as the internal research system to be used in defining major problem areas within the organization itself. (If I smile more than a little bit upon reading about a research system that helps to identify research topic

priorities, disseminates information about them to the research community, and conducts small group meetings on specific topics ... it is only that smile of recognition of seeing an analogue to some of the systems that the Marketing Science Institute has built in its special research niche over the past five years!)

Having mentioned a number of very positive things about the objectives and structure of CREB as described by the authors, what then are my questions about all this? The most significant one -- and the one which I wish the authors had discussed -- is "How does it work?" I would like a much better sense of what the experience with the system has been over the past few years -- particularly in terms of the user organizations (other governmental departments, etc.). Also, what about the assessment of the viability of this kind of system from the viewpoint of the Ministry (Consumer and Corporate Affairs Canada), and from the viewpoint of the CREB itself at both the management and the research level? Also, what is the view of the outside research community about both its actual opportunities to become involved in research and the ways in which that research has been used?

Obviously, a host of additional related questions can be raised. Again, I think that the audience can -- almost responsively in chorus -- amplify the list of sub-questions under each of the above questions.

#### CLOSING COMMENTS

As attentive listeners have no doubt recognized, I have already incorporated a number of general comments and observations in the framework of talking about each of the papers. However, I would like to add three more observations in closing -- observations which perhaps will stimulate further discussion of the topics treated today.

The first derives from the work of the TARP research and concerns the identification of problem areas and dissatisfactions of "greatest need" for action. In terms of developing priorities for public policy action (or even business action), we must be able to come up with better approaches to assessing the criteria for need and the criteria for resource allocation. All of us realize that several factors enter into any such set of considerations -- e.g., the incidence of the problem, the salience of the problem, the character of the consequences of not resolving the problem, and even the ease/difficulty of developing effective resolutions. At some future conference I would be very interested in seeing how government and business organizations are contending with this kind of problem -- given good data (such as from the TARP study) on certain dimensions of the problem.

A second comment is on the importance of learning more about the value systems that are prevalent in the public sector -- especially regarding the role of research. Do people in the public sector see research on consumer issues as an influence on public policy, as a determinant of public policy, or simply as a convenient support (when it is convenient) for their own predetermined perspectives? At future conferences, I would like to see more on this topic -- and more research on actual clinical decision-making processes on consumer issues in the public sector.

Third, I am intrigued by the Evans (et al) statement that there has been a shift in Canadian policy to put more emphasis on actions "designed to 'improve' the consumer's ability to operate effectively" within the marketplace. Even if one is engaged in regulation and intervention -- but certainly if one is trying to improve consumer ability to operate effectively within the marketplace -- one must know more about actual consumer

behaviors in the marketplace. (Note the plural because I certainly believe in segmentation with respect to this area!)

A thread that runs through all three papers -- perhaps not surprising at ACR -- is that we must know more about consumer behavior in the marketplace and we must know more about research usage behavior (and attitudes) in the public policy sector. With regard to both, we need to have more interest in consumer research from the public sector itself -- for policy formulation in the short term and in the longer term.



Ivan L. Preston, University of Wisconsin

### Abstract

This is a discussion based on the following three papers:

John L. Evans, C. Dennis Anderson, and L.G. McCabe, "Consumer Research Inputs Into Public Policy Decision Making: The Role of Canada's Consumer Research and Evaluation Branch."

Marc A. Grainer, Kathleen A. McEvoy, and Donald W. King, "Consumer Problems and Complaints: A National View."

Steven E. Permut, "Research on Consumer Information: Public Sector Perspectives."

The papers are evaluated and compared; the general topic area they encompass is discussed.

### Introduction

The thread that links these papers is the use of consumer research in the identification and solution of consumer problems by governmental regulatory units and/or marketplace participants who are involved with the problems.

The Evans *et al* paper is focussed most precisely on this topic; it discusses the development and operation of a Canadian unit called the Consumer Research and Evaluation Branch. Creation of the unit has provided the bureaucratic resources and procedures for identifying problems and either conducting research or arranging for the services of outside researchers. The scope of development described in the paper suggests that the unit may be more advanced in some respects than any comparable unit developed in the United States.

The Grainer *et al* paper is an example of the kind of research that is relevant to, and might be commissioned by, a government or non-government organization that wishes to identify and solve consumer problems. It reports a survey of types of problems that consumers experience, the damages they sustain therefrom, what they do in response, and how satisfied they are with the outcome. Identification of problems in this way is one of the natural starting points for those dealing with consumer affairs.

The Permut paper inquires into the attitudes toward consumer research held by officials of a variety of public and non-profit organizations concerned with consumer problems; it asks what research journals and what kinds of outside research advisors they find useful. The emphasis is on research done principally for researchers' own reasons and so typically having no explicitly-intended application to the specific problems which these officials must confront. Not surprisingly, attitudes toward such research are not highly favorable.

### The Evans *et al* Paper

This is not a research paper in the typical sense, so there is no point in evaluating it as such. Rather, it is a summary of the decisions and actions taken by the Canadian government in creating the Consumer Research and Evaluation Branch of its Consumer and Corporate Affairs Canada department. The description is clear

and thorough, although I would like to mention later some areas where even more explanatory material might usefully have been provided.

The development of the Branch is described first within the mandate of Consumer and Corporate Affairs Canada to create programs to protect and promote the interests of the consumer. The mandate included a specific authorization to conduct research, and the authors do a nice job of explaining developments that have strengthened the perceived need for research in recent years.

The paper then specifies in detail the mandate of the Consumer Research and Evaluation Branch, which is to identify and analyze problems and to develop, implement, and evaluate plans for remedial action. I am puzzled by the statement that the Branch actually creates and carries out remedial action; would not the role of a research unit be restricted to providing information for other people to use in implementation? Perhaps the authors have inadvertently misstated this point, or perhaps things are done differently in Canada.

Most interesting to this reviewer was the section, "The Conduct of Consumer Research," which describes the Branch's resident researchers along with its efforts to facilitate outside consumer research and to keep in touch with outside researchers. The authors describe the resident group of fifteen researchers as "small," but this reviewer does not know of that large a group of researchers in any U.S. agency devoted to consumer affairs. It may be that the Canadians are showing foresighted leadership in taking research decisions away from the lawyers and giving them to researchers. They seem also to be doing an outstanding job with the newsletters, seminars, and other means they use to inform and interact with outside researchers.

### The Content of Responses to Consumer Problems

For the most part the paper's topic is bureaucratic organization rather than consumer problems per se. Since some readers may be interested more in the content than the structure of bureaucratic responses to problems, it is appropriate that certain assumptions about the nature of consumer problems are discussed. Principally, a comparison is noted between legislative-regulatory solutions and voluntary self-adjustment solutions made by the participants in the marketplace.

The suggestion is made that public policy research ought to examine what kinds of problems should be subjected to each of these types of solutions. One might also study what kinds of adjustments members of the marketplace might reasonably be expected to make voluntarily, or refuse to make.

The authors observe that prevailing pressures nowadays are supporting the seeking of voluntary solutions rather than solutions imposed by legislators and regulators. This reviewer would have enjoyed seeing much more discussion on that topic. The paper also states that prevailing pressures operate against a long-term research perspective, as against seeking short-term solutions to specific narrow issues. Again, further discussion would have been valuable.

The paper is very useful for ACR readers in providing a summary of new developments in the interaction between consumer researchers and consumer affairs officials in government.

#### The Grainer et al Paper

The greatest strength of this study of the nature of consumer complaints lies in its contribution to correcting what the authors refer to as the "rhetoric of industry" on the topic. Industry, they say, represents consumer complaints as being typically few and not serious, and as being almost always resolved to the consumer's full satisfaction. The data show otherwise.

The study thus provides not only a much different picture of the nature of consumer complaints, but also illustrates nicely the need for research that underlies the development of the Canadian research unit described in the Evans paper.

There are both strengths and weaknesses in the fact that the study is closely related in topic to the earlier Indiana University and Center for the Study of Responsive Law studies. The weakness is that the essential findings alluded to above were already revealed in the earlier studies. The strength is that the present study used different, and in some respects better, methods of measurement. Attractively, the results correspond to those of the other studies sufficiently to support a belief in the validity of all.

#### Methodology

Attention to methodology, particularly with respect to differences among the three studies, takes up a considerable portion of the paper. Some readers may feel it is too much; others may appreciate the way in which such detail allows them to evaluate and compare.

Some of the emphasis on methodology is apparently due to the authors' desire to demonstrate the superiority of their method, for which, owing primarily to their employment of a representative national sample, they make a persuasive case. However, they perhaps do not acknowledge sufficiently the drawbacks of a method that obtained reports largely restricted to problems significant enough that consumers would recall them without being aided by a listing of products and services (probably explaining why automobile and appliance problems, mostly concerning repairs, were cited so often). Although this method is probably not inferior to the others, it is not obviously superior, either.

A methodological decision that surely seems regrettable is that we are given no demographic breakdowns, thus no indication of whether those households with different types of problems or different levels of complaint resolution, etc., may have consisted of different types of people. Nor are we given the cross-tabulations that would show whether different products or services produced different complaint profiles.

On the whole, however, the data are valuable and the study is an excellent example of the type of contribution research can make in the effort to identify and solve consumer problems.

#### The Permut Paper

While the other two papers have shown us a governmental effort to obtain research, and an example of the type of research that might be obtained, the Permut paper examines the current use of existing research on consumer information disclosure. The subjects were officials of

government and nonprofit organizations concerned with information disclosure and other consumer problems; they were asked about their use of various journals and various types of consultants, and about the extent to which they feel research contributes to policymaking.

The thrust of the inquiry was toward existing research, which means research done for its authors' own purposes and without being intended necessarily to apply directly to any current problem in the "real world." Most members of ACR probably would anticipate that policymakers would find such research to be overly abstract and theoretical, and so of doubtful value. The obtained responses confirm that expectation solidly, and this finding is the essence of the study's contribution.

Whether or not researchers should regret these findings is a question that obviously follows. We might sympathize with the failure to please policymakers, or alternatively we might insist that researchers have the right to choose topics to suit themselves, not others, and that policymakers should commission their own research aimed at their own problems. Such discussion, however, extends beyond the scope of the present study, which reports simply what the policymakers said.

#### A Problem of Interpretation

The study produces some difficulties in interpreting what the policymakers meant by what they said. Of greatest concern to this reviewer is the question of how well equipped such persons are to read and utilize existing research. The critical research question (see Table 2) begins, "Given your own experience and/or knowledge of research dealing with consumer information disclosure issues..." Well, what actually is the experience of such people about research? Furthermore, whose experience are we talking about, anyway, because there is excellent reason to believe that such questionnaires often are filled out by subordinates of the addressee?

The present study, therefore, might reasonably have included questions about how much the responding party ("if not addressee, please so state") actually knew about consumer research--was he trained in it, had he published any research himself, etc.? My own opinion, in lieu of such information, is that the questionnaire's respondents probably were rather low on such knowledge. Further, I would surmise that the lower the research sophistication, the more negative the attitudes toward research. Thus a possible alternate explanation of those reported negative attitudes.

A number of other methodological questions, involving such things as sampling, statistical analysis, choice of question, and interpretation of findings, might be raised about this study. The results, however, are essentially valuable for the important issues they raise.

#### Researchers and Regulators: The Expanding Interface

There is good reason to believe that the interaction of consumer researchers with regulators and other policymakers will continue to increase in the future as it has in the recent past. Research papers such as those discussed here are therefore most welcome as a means of contributing to the success of such interaction.

What this reviewer would like to see as a follow-up to such contributions would be a detailed study of the participation of research and researchers in regulatory affairs. I have recorded briefly elsewhere some aspects of my own participation in Federal Trade Commission cases in which deceptive advertising is charged, with emphasis on the problems involved when researchers work with lawyers. However, a definitive look at all

features of such interaction has yet to be written.

Some of the inherent problems are alluded to in the papers examined above. For example, the Grainer paper discusses methodological differences among three problems, which brings up the problem of persons untrained in research attempting to make legal judgments about research content. I have personally found FTC lawyers to be bright and quick in grasping explanations of such matters, but have also noted that such a learning process is typically piecemeal, informal, and incomplete.

#### The Need For Training

What needs to be created, therefore, are training sessions that can do the job more formally and thoroughly. The trainees should include not only those who write and prosecute rules, but also judges, both within the regulatory agencies and on the appellate courts. Staff attorneys who deal with researchers have considerable opportunity in working sessions to acquire detailed information on the background of relevant questions, but judges interact with researchers only within the severe confines of a hearing, which restricts discussion to the specific facts of the case at hand. Yet it is the judges who make the ultimate decisions about the probative value, as they call it, of research studies.

Nor should the training sessions be restricted to the regulators. Researchers need to learn about the law, and about the political realities behind the law, just as badly as lawyers need to learn about research (see Permut's survey questions on this point). One might argue, in fact, that the researcher's need for orientation is even greater than the lawyer's, because the interaction takes place on the lawyer's "turf." It is his game that is being played.

#### Knowing the Differing "World Views"

Training of lawyers and researchers to understand each other should include not merely techniques, but also, as Permut discusses, some understanding about the differing "world views" of the two occupations. When Scott Ward and Dan Wackman testified about children's advertising before the FTC in 1971, Commissioner Mary Gardiner Jones illustrated these differing views by asking for an action recommendation. The researchers answered, "More research is needed before we can say"; the Commissioner said, "I have to act now."

Ms. Jones implied that researchers, if they were to be helpful to her, would have to be willing to make recommendations under conditions of greater uncertainty than they were accustomed to experiencing. Whether researchers should do that, and how they might do it, might itself be the topic of a useful research study.

Elsewhere, Permut asserts that the participation of the researcher in the activities of regulators and policymakers should be the researcher's responsibility. This presumably is because researchers via their methodology have greater access to the facts of a case or problem, and therefore should accept the obligation to assure that the chosen solution is properly based on those facts. It is a sensible analysis, but I would like also to point out that there is no way a researcher can participate in decision-making without the consent of the regulators or policymakers. Only the latter can make the decision to utilize existing research findings and/or commission appropriate new research.

It might follow, therefore, that the researcher's main responsibility is better defined as being that of convincing the regulators and policymakers that research is their responsibility. In any event, the question of responsibility is an area that merits more study than it has been given thus far. One aspect of such study

might be to examine the shift from minimal to considerable interaction with behavioral researchers that has occurred at the FTC over the past seven years.

#### Education Versus Regulation

Another area that deserves further study is that of the relative attractions of consumer education and regulation as the best choice for responding to consumer problems. This topic is mentioned several times in the Evans paper; the authors report that there are reasons today for seeking non-regulatory solutions, of which consumer education is an important kind.

The issue of regulation versus education is fascinating because of the conflicting assumptions it involves about the basic nature of the human being. Proponents of education support the assumption that the consumer is educable, not only in having the natural ability to become educated but also in having the time and money it takes, and perhaps most importantly, the inclination. They also assume that education is enough to solve consumer problems, because the consumer upon learning the relevant facts will act to eliminate any problem.

Proponents of regulation, on the other hand, assume either that the consumer is not educable on a given point or else assure that even if maximally educated he will be unable to eliminate the problem unless alterations in the conditions of the marketplace are imposed by the regulators.

These conflicting views produce still another area in which research can contribute; it can determine in given cases whether consumer education is sufficient or whether regulation is necessary, thereby having an important impact on the direction taken by public policy.

## A STUDY OF PUBLIC POLICY IMPACT ON CONSUMER DECISION-MAKING

Philip C. Burger and Alladi Venkatesh  
State University of New York at Binghamton

### Abstract

This report describes how a consumer decision-making model was created and tested in research sponsored by the Federal Communications Commission. The FCC was considering a series of policy changes with regard to equipment and rules of operation. Discriminant and regression statistical results were combined into a simulation model which projected the impact of the proposed policy changes.

### Introduction

Traditionally, marketers have focused their attention on the impact of corporate policies upon consumer decision-making and choice behavior. The policies of the government were a given condition external to the consumer decision-making framework. Political scientists and economists, however, have always been interested in the politics and economics of regulation at the aggregate rather than the individual level.

This study was commissioned by the Federal Communications Commission to explore the impact of federal policy making on consumer decisions. For our purpose, 'policy' can be defined as a series of rules affecting the availability of products and services with the express intention of achieving certain objectives. In the case of the government, such objectives have a broad social content and in the case of a corporation they are usually profit oriented. Both government and corporate policies can result in the creation and elimination of products which in turn have a direct impact upon consumers.

The Federal Communications Commission originally commissioned this particular study to determine which of several policies might be implemented to improve citizen's band radio service. Citizen's band radio had grown from under 40,000 users in 1959 to 20 million users in 1977. The frequencies available to CB users created certain limitations for personal radio communications. Being aware of this problem, the FCC was developing a number of alternative proposals. At the same time, it was getting concerned with possible consumer reactions to these policy proposals. There was, obviously, a growing desire within the FCC that any new policy should maximize the satisfaction of as many CB users as possible.

### Background

Typically, studies on public policy issues relating to marketing have been in the areas of consumerism, deceptive advertising, legal restraint, products labeling, and such other issues involving consumer protection (Aaker and Day, 1978). Commenting on the role of consumer research, Sheth (1974) noted "the bulk of research in consumer behavior for public policy is likely to be problem oriented, and most of the problems are likely to come from the negative side effects of mass marketing and mass consumption - for example, mass media effect on citizen's value." Recently, however, some authors have noted that substantive research is needed on consumer decision-making and choice behavior as inputs to public policy making or to evaluate its impact (Wilkie, 1976, Jacoby, 1977). There are very few systematic empirical or theoretical studies at the micro consumer behavior level in the marketing literature.

What little work there is, has been done evaluating FTC policies on deceptive advertising, children's commercials, etc.\*

Our study marks a serious departure from the existing public policy-marketing interface issues and examines consumer attitudes and usage patterns of a highly regulated product, CB radios. In evaluating the impact of policy making upon consumer attitudes and decisions, a segmentation design was developed based on CB user/non-user attitudes, psychographics and demographics. In addition, the research analyzed the recent adoption behavior of CB radio users and extrapolated future adoption possibilities resulting from key alternative policy proposals. Thus the study, while indicating that some generalizations could be made about consumer behavior in response to policy making, also pointed the way toward picking a more optimal set of policy alternatives.

### The Policy-Consumer Impact Model

Figure 1 shows a simplified model on which most marketing studies are based. The process underlying the model suggests that a given federal agency creates a set of rules concerning a relevant marketing issue such as advertising, product safety, etc., which are transmitted to industry (arrow 3 in Figure 1). Industry then uses such rules as inputs into its product and promotion policies which, in turn, impact customers (arrow 4). There is a feedback loop from the customer both back to the industry and the agency. An example of this is the use of seat belts. In spite of the auto safety implications of the usage of seat belts, car drivers have not been using them. Consequently, the government is trying to enforce this, not directly, but by insisting that manufacturers build an automatic seat belt mechanism in the cars.

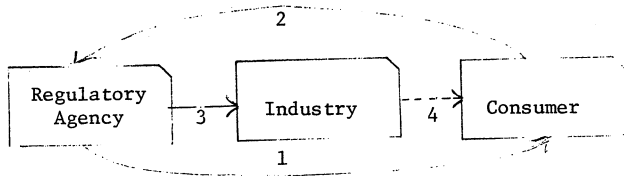
A more complicated set of relationships is suggested by the arrows 1, 2 and 3 in Figure 1. This system of effects may be described as the general paradigm for the present study. The sequence of steps are simply that the regulatory agency defines a policy which is transmitted to consumers through mass media or any other means (arrow 1). The consumer reaction in terms of letters, complaints and other types of communication is relayed back to the regulatory agency which then modifies its position, if necessary, and suggests a set of technical rules and regulations to industry (arrow 2). Based on such directives from the government, the industry makes its own product policies and the cycle continues (arrow 3 and 4). This paradigm clearly illustrates what happened in the case of the CB radio beginning in 1959. The early ruling concerning CB radios set aside the 27 MHz band for use of private citizens using 4 watt (low power) radios for their own purposes. These short range radios achieved very limited adoption. The regulation generated very little reaction among consumers initially. However, in 1973 when the "truckers strike" occurred as a result of the Arab Oil Embargo, the FCC was bombarded by license applications for the new radios. A great demand for the radios existed and there was a small supply. The industry responded by building many radios and im-

\* For a very interesting theoretical model designed to evaluate welfare effects of regulation at micro consumer level, using Lancaster's approach, see Colantoni, et al. (1976).

porting many from Japan. Large numbers of applications were received by the FCC and a large number of complaints about channel crowding were also received (Figure 1, arrow 2). The FCC responded (arrow 3) by adding 17 additional channels to the existing radio configuration.

FIGURE 1

INITIAL POLICY - CONSUMER PARADIGM

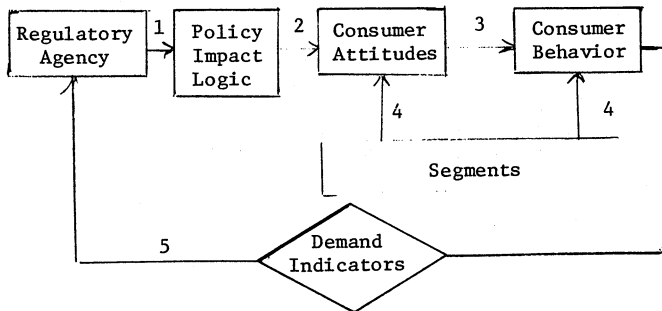


Consumers responded again to the FCC with complaints about the telephonic quality of the radio and the FCC then began considering new alternatives which would change the configuration and the price of the radio service. In addition, it became clear that a certain culture was developing around the citizen's band radio. This culture manifested itself in usage behavior, special language, and other unforeseen uses for the radio itself which included locating police radars, providing and enhancing security of individuals, boating uses, and other legal and semi-legal uses.

It became clear in discussions with the FCC that a more refined approach to the policy impact on consumer decision-making was required. Therefore, the relationships shown in Figure 2 were introduced as the basis for the research. This new paradigm can be viewed as a circular flow beginning with the regulatory agency. The regulatory agency promulgates rules and technical specifications for products. These specifications impact on the cognitive structure of the consumers, the logic of which is created by the interfacing of the technical aspects of the product with the perceptual aspects of the products and service. In terms of CB radios, consumers perceive CB radios as a certain bundle of characteristics a subset of which is influenced by the policies of the FCC whose impact is also clearly perceived. One resulting hypothesis is that the policy impact alters consumer's beliefs-attitudes and which in turn affects purchase and usage behaviors.

FIGURE 2

EMERGENT POLICY CONSUMER PARADIGM



The segmentation aspect of the model (represented by arrow 4) posits that different consumer groups react in different ways which are unique to the segment. Arrow 5 in Figure 2 indicates that consumer behavior responses can be integrated into an indicator system which can be monitored by regulatory agencies. Thus, the model can be viewed as a stimulus-response model where, the regulatory agency creates a stimulus and the consumers respond over a period of time. The regulatory agency must then interpret the demand indicators to ascertain whether the policy contributes to consumer satisfaction or not.

Sampling Design

The sampling was done on a national basis. Two types of primary sampling units (psu's) were identified, metropolitan areas and the others. All the psu's were grouped into 65 strata. The secondary sampling units (ssu's) were identified as all areas with a population of approximately 2500 (e.g., a block or two in a major city; a census tract in some rural areas). Systematic random procedures were used to select 100 ssu's. Finally, households were selected from the ssu's. In each household, the interview was conducted with the CB radio user, and for nonuser households, the key male member of the household was questioned.

In the first stage, about 863 users and nonusers were personally interviewed. This was followed by 441 interviews with CB user households which were identified through telephone screening. The final sample included 1304 respondents of which 674 were nonusers and the remainder were users. Because of the disproportionate representation of the users in the sample, a weighting scheme was devised to achieve representative estimations in the analysis of sample data.

The data was collected as shown in Table 1. The questionnaire consisted of five major sections relating to CB ownership, usage, and attitudes. The first section included detailed information on the ownership of CB radios, as to when they were purchased, what price was paid, and the number of channels the radio had. The second major section of the data was the usage patterns of the CB radios. This data included time of day, number of hours per week, terrain of use, and purpose of use. A third major section of the questionnaire involved a 19 item semantic-differential scale measuring beliefs and attitudes toward various aspects of CB radios. The list of attributes was derived from detailed focused group interviews with both users and nonusers of CB radios.

TABLE 1

DATA COLLECTED

1. Ownership of CB radios.
2. Usage patterns of CB radio
3. Semantic Differential Beliefs and Attitudes (19 items)
4. Psychographic data
5. Demographic data

The fourth major section of the questionnaire involved 35 questions which were derived from the psychographic literature (Wells and Tigert, 1971). The questions devised by previous researchers were used with almost no modification. Of the fourteen life style factors proposed by Wells and Tigert, 12 were hypothesized to be related to CB radio usage.

The last major section included demographic information including age, income, occupation, residence, number of cars owned, education, and others. The questionnaire took about an hour to complete.

The Results

The data analysis used different techniques: (1) basic descriptive statistics, (2) cross-tabulation, (3) segmentation analysis, (4) discriminant analysis, and (5) regression. At the end, the statistical information was cast into a computer simulation program to test effects of various policy impacts on the dependent variables.

The dependent variables in the study, which were the basis for the demand indicators are: (1) projected numbers of users and nonusers by segment, (2) the pattern

TABLE 2

## FACTOR ANALYSIS OF PSYCHOGRAPHICS

<u>Factor 1</u>	<u>Factor 2*</u>	<u>Factor 3*</u>	<u>Factor 4</u>	<u>Factor 5*</u>	<u>Factor 6</u>
I usually watch the ads for announcements of sales.	I like to watch or listen to football or baseball games  I usually read the sports page in the newspaper.  I thoroughly enjoy conversations about sports.	I often repair my own car.	I engage in many intellectual activities in my spare time.	I am an active member of more than one service organization.  I do volunteer work for a hospital or service organization on a fairly regular basis.	No variable had a factor loading of .40 or greater
Variance: 10.3%	8.2%	6.8%	5.8%	5.4%	4.8%
<u>Factor 7</u>	<u>Factor 8*</u>	<u>Factor 9*</u>	<u>Factor 10</u>	<u>Factor 11</u>	
I enjoy activities such as going to museums, plays and concerts.  I wish I had more time to do artistic things like dancing, writing or painting.	I spend more of my budget on my clothes than other people in my income class.  An important part of my life and activities is dressing smartly.	I spend a lot of time talking with my friends about local events and people.	I enjoy my privacy  I like to withdraw from others for extended periods of time.	I sometimes think when people have a misfortune, they only got what they deserved.	
Variance: 4.4%	3.3%	3.2%	3.1%	3.0%	

\* Significant factor as determined by analysis with user/nonuser groups

Total Variance: 58.3%

of usage, (3) the number of hours per year of usage by segment, and (4) attitudes in terms of satisfaction/dissatisfaction with elements of the radio service.

The segmentation analysis was undertaken using factor analysis as a prime technique. First, the psychographic variables were factored to see if the loadings and factor patterns were consistent with those found by Wells and Tigert. The results appeared to confirm their findings. The emergent factor patterns are shown in Table 2 with relevant segments identified by asterisks. Only those psychographic variables which discriminated best the users from nonusers in cross tabulation analysis and which also loaded highly on the factors were considered important in the segmentation analysis. Six meaningful segments emerged. The first one was the indoor sports fan. The indoor sports fan watches television, reads sports magazines, and values entertainment. This was a prime segment of the CB market. This segment accounted for 26.6% of the U.S. population and over 20% of this particular group owns CB radios. The second major segment was the car repair buff. These people constituted approximately 18.8% of the U.S. population and almost 31% of them owned CB radios.

The third segment was the person who belongs to the service organization. These people were the members of the CB radio clubs, ambulance services, volunteer fireman, and such. These people comprised approximately 6% of the U.S. population but over 31% of them own CB radios. A fourth major segment was the fashion conscious businessman. This person values his or her appearance and spends a great deal of time traveling. This segment accounted for 6.5% of the population and almost 18% of this group owned CB radios.

Segment 5 is called the residual under-30 group consisting of those individuals who were in the age range from

16 years to 30 years, but did not fall in the previous four segment categories. They accounted for approximately 11% of the population and 16% of them owned CB radios. The sixth and last segment includes the residual middle income person. These people are all those who did not fall in the previous five segments but who were classified as middle income (\$16,000-\$25,000). They comprised about 11% of the U.S. population and almost 12% of them have CB radios. The null segment or the members of no segment accounted for approximately 18% of the U.S. population and 6% of them had CB radios. Additional segments were not considered important because of the low marginal information obtained by adding more segments.

After the segments were identified, all of the data analysis was performed on a within segment basis. The next major step was to see if attitudes and perceptions of radio service discriminated between users and nonusers in each segment. It was hypothesized, based on the results of the focus group interviews, that belief and attitude variables would tend to discriminate between users and nonusers.

A discriminant analysis was performed and the results are shown in Table 3. In the segment containing service organization members, a nonsignificant user/nonuser discrimination was discovered. Indoors sports fans valued being entertained. The car repair buffs valued entertainment, helping others and getting help. Also, the aggregate preference score which was a simple summation of all the attitudes values appeared is a discriminant for the car buffs.

The fashion conscious business segment was male, and valued making friends and being entertained. The residual middle income group valued entertainment and helping others. The necessity for boating communications was

evaluated negatively. Thus, it was clear from the data analysis that attitudes were very powerful discriminators between users and nonusers. It was also clear from the data analysis that other variables such as demographics and car ownership, etc., were not as powerful discriminators.

TABLE 3

VARIABLE DISCRIMINATING USERS FROM NONUSERS

Indoor Sports Fan

	<u>Sign</u>
Being Entertained	+
Helping Others	+
Getting Directions	+
Getting Help	+
Learning Road Conditions	-

R<sup>2</sup>: 25.0                      p < .0001

Car Repair Buff

	<u>Sign</u>
Aggregate Preference	+
Being Entertained	-
Getting Help	-
Helping Others	-
Getting Directions	-

R<sup>2</sup>: 16.0                      p < .0001

Service Organization Member

None

Fashion Conscious Businessman

	<u>Sign</u>
Sex	Male
Making Friends	+
Being Entertained	+

R<sup>2</sup>: 64.0                      p < .001

Residual Under 30

	<u>Sign</u>
Sex	Male
Aggregate Preference	-
Boating Communications	-
Helping Others	+
Getting Time	+
Learning Road Conditions	+

R<sup>2</sup>: 21.3                      p < .0001

Residual Middle-Income

	<u>Sign</u>
Sex	Male
Being Entertained	+
Helping Others	+
Boating Communications	-

R<sup>2</sup>: 21.0                      p < .0001

Regression analysis was also performed to predict the various kinds of CB usage behavior. As a first step, all of the behaviors were factor analyzed and found to group into three basic factors. One of these was, time of day, namely 2 P.M. to 10 P.M. daily. Another was use in the suburbs and small cities while the last was use in flatlands and low hills. These three factors became dependent variables of usage. For each segment, a series of regressions using attitudes as the independent variable were performed to see if attitude structure predicted usage pattern. The results are shown in Table 4. The R<sup>2</sup> range from 3% to 34% indicating that in some types of usage a very low predictive power but others have quite high predictive power. Overall, the prediction of usage pattern was much weaker than discriminating users/nonusers. It is important to note that, in general, the results show a tradeoff between telephonic needs and entertainment. The indoor sports fans wished to be entertained while other groups tended to value telephonic communication.

Using the Results to Test Policy Alternatives

A computer simulation program was created using the results from the regression and discriminant analyses. The program which was written in the FORTRAN language, contained all of the regression/discriminant b coefficients for each segment. The program used the raw data from each respondent as input, classified each respondent as a user or nonuser of the proposed CB service, determined the rate of usage by category and computed some aggregate sample statistics. The input data was weighted to reflect a sample representative of the U.S.

The key logic of the simulation required that changes in perception of the proposed radio service be related to the engineering aspects of the proposal. The

TABLE 4

VARIABLES PREDICTING BEHAVIOR

	<u>Sign</u>	<u>Indoor Sports Fan</u>	<u>Dependent Variable</u>
<u>Variables</u>			
Being Entertained	+		2 p.m. - 10 p.m.
Using Latest Device	+		
		R <sup>2</sup> : 9.5	
Being Entertained	+		Use in suburbs
Hunting/Fishing	-		
		R <sup>2</sup> : 3.6	
Being Entertained	+		Use in flatlands
Hunting/Fishing	-		
		R <sup>2</sup> : 4.9	
		<u>Car Repair Buff</u>	
<u>Variables</u>			
Way to Make Phone Calls	+		2 p.m. - 10 p.m.
Boating Communications	-		
Getting Time	-		
Using Latest Device	+		
		R <sup>2</sup> : 9.0	

Car Repair Buff (Cont'd)

<u>Variables</u>	<u>Sign</u>	<u>Dependent Variable</u>
Hunting/Fishing	-	Use in suburbs
Finding Radar	-	
Talking to Home	+	
Way to Make Phone Calls	+	

R<sup>2</sup>: 7.4

Feeling Part of "In-Group"	+	Use in flatlands
Talking to Work	+	
Hunting/Fishing	-	

R<sup>2</sup>: 4.6

Service Organization Members

<u>Variables</u>	<u>Sign</u>	<u>Dependent Variable</u>
Feeling Part of "In-Group"	-	2 p.m. - 10 p.m.
Making Friends	+	
Using Latest Device	+	
Getting Time	-	
Getting Help	-	

R<sup>2</sup>: 27.0

Learning Road Conditions	-	Use in suburbs
--------------------------	---	----------------

R<sup>2</sup>: 3.2

Talking to Work	+	Use in flatlands
-----------------	---	------------------

R<sup>2</sup>: 4.7

Fashion Conscious Businessman

<u>Variables</u>	<u>Sign</u>	<u>Dependent Variable</u>
Talking to Work	+	2 p.m. - 10 p.m.
Boating Communications	-	
Getting Help	+	

R<sup>2</sup>: 34.3

Talking to Work	+	Use in suburbs
Boating Communications	-	
Hunting/Fishing	-	

R<sup>2</sup>: 25.5

Boating Communications	-	Use in flatlands
Making Friends	-	

R<sup>2</sup>: 33.4

Residual Under 30

<u>Variables</u>	<u>Sign</u>	<u>Dependent Variable</u>
Getting Time	-	2 p.m. - 10 p.m.
Talking to Home	+	
Hunting/Fishing	-	
Using Latest Devices	+	

R<sup>2</sup>: 11.4

Hunting/Fishing	-	Use in suburbs
Getting Time	-	
Talking to Home	+	
Being Entertained	+	

R<sup>2</sup>: 13.5

Residual Under 30 (Cont'd)

<u>Variables</u>	<u>Sign</u>	<u>Dependent Variable</u>
Hunting/Fishing	-	Use in flatlands
Getting Time	-	
Being Entertained	+	

R<sup>2</sup>: 10.3

Residual Middle-Income

<u>Variables</u>	<u>Sign</u>	<u>Dependent Variable</u>
Way to Make Phone Calls	+	2 p.m. - 10 p.m.
Boating Communications	-	

R<sup>2</sup>: 5.4

Way to Make Phone Calls	+	Use in suburbs
Finding Radar	-	

R<sup>2</sup>: 3.1

None		Use in flatlands
------	--	------------------

program simulated changes in beliefs and attitudes of each individual respondent and computed adoption and usage figures based on the changes. Discussion with FCC personnel revealed each proposal primarily affected two aspects of radio perception: (1) the telephonic quality (the ability to transmit information over substantial distances with little or no interference) and (2) the entertainment value (being able to listen to other persons chat). The technical proposals served to enhance telephonic quality while reducing entertainment value. Price was also included in the simulation. The gross relation between price and adoption rate was investigated and found to be a decreasing function of price with an elasticity of  $b = -1.4$ .

Conclusions

The FCC was considering two primary proposals. The first of these proposals would be to add a new CB radio service at 220 MHz. This service would allow many more channels than are currently available. The net effect of this would be to decrease the entertainment value of the CB radio by diffusing users over a larger number of channels. However, on the positive side, the service would become much more telephonic in its ability to handle communications. A second major policy proposal was to open up a very high frequency band in the 900 MHz range which would increase the telephonic aspect of the CB radio quite substantially and would further reduce its entertainment value.

Given these logical specification, it can be seen that entertainment and phone communications appear in both the discrimination and the usage pattern predictions. Thus in the simulation, the question becomes whether the increase in telephonic communication would outweigh the decrease in entertainment value. The overall evaluations of the FCC Proposal showed that there was a small yet significant market for the new services.

References

1. Aaker, David A. and George S. Day, "Consumerism: Search for the Consumer Interest," New York: The Free Press, 1978 (Third Edition).



References (Cont'd)

2. Bloom, Paul N. and Nikhitesh Dholakia. "Marketing Behavior and Public Policy: Some Unexplored Territory," Journal of Marketing, 37, 4 (Oct. 1973), 67.
3. Colantoni, Claude S., Otto A. Davis and Malati Swaminathan. "Imperfect Consumers and Welfare Comparisons of Policies Concerning Information and Regulation," The Bell Journal of Economics, Autumn, 1976, 7, 2.
4. Dyer, Robert F. and Terence A. Shimp. "Enhancing the Role of Market Research in Public Decision-Making," Journal of Marketing, 41, 1, January 1977, 63-67.
5. Jacoby, Jacob. "Information Load and Decision Quality: Some Contested Issues," Journal of Marketing Research, 14, November 1977, 569-573.
6. Jones, Mary Gardiner. "The FTC's Need for Social Science Research." (Address before the Second Annual Conference, Association for Consumer Research, College Park Maryland, September 1971).
7. Richardson, Lee. Consumers in the Federal Decision-Making Process, California Management Review, 16, 2, Winter 1973, 79-84.
8. Ross, Ivan. "Applications of Consumer Information Public Policy Decisions," in Jagdish N. Sheth and Peter L. Wright eds., Marketing Analysis for Societal Problems, Champaign, Ill, University of Illinois Press, 1974.
9. Sheth, Jagdish. "The Next Decade of Buyer Behavior Theory and Research" in J. Sheth (ed) Models of Buyer Behavior: Conceptual, Quantitative and Empirical, New York: Harper and Row, Publishers, 1974 (ch. 21).
10. Summers, John O. "Less Information is Better?," Journal of Marketing Research, 11, November 1974, 467-468.
11. Wells, William D. and Douglas J. Tigert. "Activities, Interests and Opinions," Journal of Advertising Research, 1971.
12. Wilkie, William L. "Consumer Information, Acquisition: Public Policy Perspectives," Association of Consumer Research Proceedings, 3, 1976, 334-340.

## A MEASURE FOR MARKET DELINEATION

Jacques C. Bourgeois, Carleton University

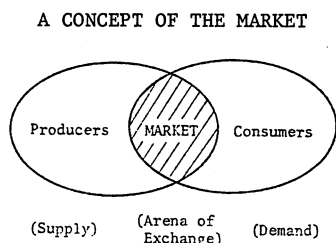
### Abstract

The present paper briefly reviews those measures used in the market delineation task. However, its main purpose lies in presenting an objective means to accomplish this task. The proposed methodology derives a statistic based on consumer derived judgemental data.

### Introduction

The concept of a market has been described in many ways, but it still remains elusive. Often specific applications have appeared in the literature, such as the beer market, the retail market, the computer market, the job market, the youth market. All of these views of a market have two things in common: 1) they include a group of consumers seeking to fulfill a need, and 2) a group of producers seeking to deliver need-satisfying products. Thus, a market may be defined as the arena of potential exchanges between consumer and producer. The word "potential" is used to imply the possibility of future markets based, for example, on unsatisfied needs or the development of new technology. A market is a set of consumers and producers with something in common. That is, they both want to trade to fulfill a need. The consumer seeks to "procure" need fulfilling products, while the producer seeks to "deliver" need fulfilling products. Thus, it could be viewed as the intersection of (or exchange arena between) a group of producers (supply) and a group of consumers (demand). Perhaps the following simple graphic model best illustrates this concept.

FIGURE 1



Others have similarly expressed this perspective of a market. For instance, Steiner (1968) defined a market as "the entire web of interrelationships between buyers, sellers, and products that is involved in exchange". Nickels (1978) defines a market as "a particular group of people and/or organizations that have wants and needs that may be partially satisfied through a marketing exchange".

### Literature Review

Several definitions have been used in identifying a market and generally these are presented as a function of the author's field of study (e.g., economics, marketing, law). It is not the purpose of this review to recapitulate these definitions. A detailed review of these definitions is contained elsewhere (Day and Shocker, 1976). The literature reviewed in this paper will be concerned with those "measures" used to operationalize these definitions. The review and format which follows,

draws from that suggested in an earlier paper (Day and Shocker, 1976). The literature has been dichotomized following two general perspectives: 1) behavioral measures and 2) perceptual measures. This division has previously been suggested by Krech and Crutchfield (1948). They suggested the use of these two perspectives as a proper means for recognizing a "group". In our context, a group is interpreted as that set of buyers and sellers in the process of exchange - a market. It is also interesting to note that a similar categorization of measures has been applied by Day and Shocker (1976) when they used a behavioral and judgemental dichotomy.

### Behavioral Measures

Bass, Pessemier and Tigert (1969) identified, through factor analysis, groups of products in terms of their similarity in usage rates. The resultant factors were groups of products considered to be substitutes or complements. This procedure has been questioned (Day and Shocker, 1976) in terms of the relevance of the criterion. That is, it would seem to be a big assumption to use synonymously similarity in purchase rates and the concept of substitutability.

The concept of cross-elasticity of demand is another behavioral measure. It is one that has had some acceptance in the courts (Dean, 1962 and Werth, 1965). The idea of product substitution has gained attention not only for consumer products but, more recently, for industrial products, due to the increasing costs of raw materials. For instance, Business Week (Sept. 14, 1974) reported the efforts of three major soft drink makers (including Coca-Cola) who made a switch from sugar to high fructose corn syrup. Business Week also examined other examples, but all served to illustrate that when a product can act as a substitute (high cross-elasticity of demand), then these products can be considered in the same market. On the other hand, Day and Shocker (1976) suggest that despite the impressive credentials of this measure, it is widely criticized and infrequently used. This would be due largely to the fact that it is based on an assumption seldomly satisfied - that is, it assumes that there is no response by one firm to a change in marketing variable(s) of another.

Brand switching measures are also behavioral measures based on conditional probabilities of purchasing a particular brand or product. Depending upon the transition patterns between products, markets can thus be defined. Two products with high transition probabilities would be considered in the same market. This approach also suffers from certain shortcomings, namely that it relies on the stability of the customer choice process and on the feasibility and reliability of using panel data.

### Perceptual or Judgemental Measures

Decision sequence analysis (Bettman, 1971, 1974; Haines, 1972, 1974) and similarly the application of utility trees in economics (Strotz, 1957, 1959) are judgemental measures used for the market delineation task. Basically, these examine the consumer's process in evaluating alternatives. The set of choices are set along a decision hierarchy - a sort of lexicographic model. This decision hierarchy is then used to partition the market. However, there exist problems with this technique as well. Much of the analysis of protocol data used in

decision analysis deals at the individual level and aggregation over a segment still poses a problem in this area of research. On the other hand, the operationalization of a utility tree has more recently met with some success (Bourgeois, Haines, Sommers, 1975).

Customer judgements in terms of their perceptions of product usage have been used to develop market boundaries (Stefflre, 1971). In this instance, an "items by uses" matrix is collected such as to yield a similarity matrix. Then, all similar products within this matrix are placed near each other, such as to have product clusters stand out. The problems here lie in determining how similar the products must be before they are clumped. In addition, the technique employed to create the respondent-generated product set remains, at this point, subjective, although Srivastava, Shocker and Day (1977) seem to have recently offered a solution to this problem. A somewhat similar approach to that of Stefflre has been suggested by Bourgeois, Haines and Sommers (1975). In this instance, respondents are asked to group brands; respondents arbitrarily determine the number of groups used, as well as the criteria for forming these groups. Based on the homogeneity of respondents' perceptual judgements, market boundaries are defined. The shortcoming here is that the respondent's "choice sets" are made independent of "intended usage" (or usage situation). Thus, it may be possible for two products to be substitutes in a particular situation, while in a situation with completely different "intended usage", we might find these same two products not so closely related.

#### A Statistic for Market Delineation

Given the above brief review, what is needed is a measure which relies on the basic definition presented at the start of this paper, while overcoming the shortcomings presented in the literature review. The following statistic is derived as such a measure. While using the above dichotomy, one may class this approach as a consumer based judgemental measure.

The problem in delineating a market lies in finding the intersection of demand and supply for need fulfilling products. That is, we must identify a group of products supplied or available to the consumer which are also perceived by these consumers as substitutes. It has been stated Moran, 1973), at the highest level of generalization, that, "to some degree, in some circumstances, almost anything can be a partial substitute for almost anything else". At the lowest level we may find, at most, two similar products, and indeed, two similar products may not even exist. Thus in fact, there is a similarity continuum, that is, a degree of similarity between any two and all products. The problem lies in identifying and measuring a product's position along this continuum.

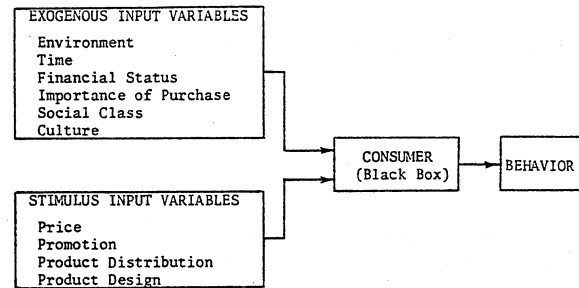
Given that a market is that arena of exchanges where supply and demand meet, an operationalization of such a definition should include the producers' outputs, as represented by a total set of offerings - a product set. On the other hand, the demand side is represented by consumers' wants for need satisfying products. Ideally, each consumer should find a product to satisfy his exact needs. This would actually involve producing one product for each individual. Obviously, this is not a practical solution, given our present technology. Thus, producers offer what they believe to be the correct set or products while consumers satisfy by choosing among the available alternatives.

In choosing among the available product set, consumers must trade-off between the various alternatives. This trading-off is affected by external (e.g., marketing

variables) and internal (e.g., perceptual distortion, perceived risk) mechanisms. The part research has typically built into their approach an operationalization and measure of these variables by, for instance, relating perceived risk to behavior or the effect of changes in certain marketing variables to product selection. It is the purpose of this approach to treat these internal and external mechanisms as part of a total system of interacting forces - a Gestalt. Thus, the focus here is not to measure each of these components but rather to measure the resultant output from these interacting mechanisms (see Figure 2). That is, given a present set of variables describing a situation at a particular point in time, what is the consumer's perception of the market? Given a set of inputs, what is being measured is simply the output from the consumer's black box.

Figure 2

#### A SIMPLIFIED INPUT-OUTPUT MODEL OF CONSUMER BEHAVIOR



Source: Adapted from The Theory of Buyer Behavior, by Howard, J.A. and J.N. Sheth, Wiley, 1969.

The task thus lies in measuring consumers' perception of product substitutes given the set of total available products. This task or research problem should be approached from two perspectives (as previously stated): a demand and a supply dimension. The supply dimension is relatively easily measured, as firms produce what they perceive and believe consumers want. The output on the supply side is clear - a product(s). The demand side is not so clear. Given the above illustrated input stimuli and interacting internal mechanisms, it is difficult to measure their output or perception of product substitutes. The difficulty lies mainly in the fact that investigators have relied largely upon surrogate measures of this product substitution phenomenon. The present technique uses a more direct route by simply asking consumers which products they perceive as substitutes. These subjective judgements are taken to be direct measures of a consumer's perception of product substitutes. They then serve as maximum likelihood estimators of the underlying parameters for the market delineation task.

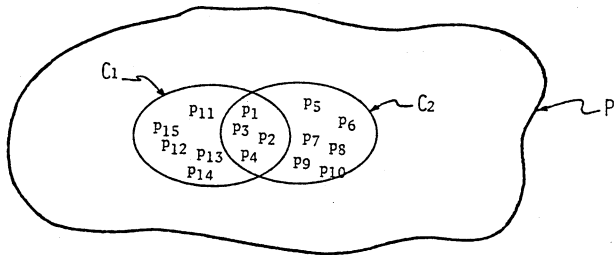
#### A Two Consumer P-Product Market

Before expanding the model to a whole population, it may be easier to show an example for the case where we would have simply two consumers and P products (see Figure 3).

The P set represents that group of products offered by all producers; these are thought, by producers, to be the total set of products satisfying a need. In fact, in this two consumer population,  $C_1$  perceives nine products ( $P_1 \dots P_4, P_{11} \dots P_{15}$ ) as substitutes, while  $C_2$  perceives ten products ( $P_1 \dots P_{10}$ ) as substitutes. In this simple case, each consumer could form separate market segments. But, if one were to define a MARKET for this situation, it would be that sub-set of products which is perceived similarly by both

FIGURE 3

THE MARKET - THE TWO CONSUMER P-PRODUCT CASE



consumers, or in set notation:

MARKET SEGMENTS:  $C_1$  and  $C_2$

MARKET:  $C_1 \cap C_2$

Some products are neither part of the market nor of market segments. This would be the set of products in:

$(C_1 \cup C_2)'$  or  $P - (C_1 \cup C_2)$

These products are not perceived as being in either the market or market segment and thus should not be produced for this particular market; if they are, they should be satisfying some "other" need. The present market is therefore defined as the set of those four available products being perceived as substitutes by the two consumers. We may say that we are 100% certain that these four products are in the same market, since both consumers of this two consumer population perceive them as such. On the other hand, products  $P_5 \dots P_{15}$  are not as clearly defined. For instance, we could say that we are only 50% certain that each of the groups of products  $P_5 \dots P_{10}$  and  $P_{11} \dots P_{15}$  forms separately a market, since these are respectively perceived as being in a market by only one of two consumers. In probabilistic terms:

$$P_{P_{11} \dots P_{15}} = \frac{\text{no. of consumers } \in (C_1 \cup C_2) - C_2}{2} \quad (1)$$

$$= 0.5$$

$$\text{OR: } P_{P_5 \dots P_{10}} = \frac{\text{no. of consumers } \in (C_1 \cup C_2) - C_1}{2} \quad (2)$$

$$= 0.5$$

A Three Consumer P-Product Market

A second example will be given to illustrate the problem. Let us suppose that our population was made up of three individuals with the same set of products as in the above example.

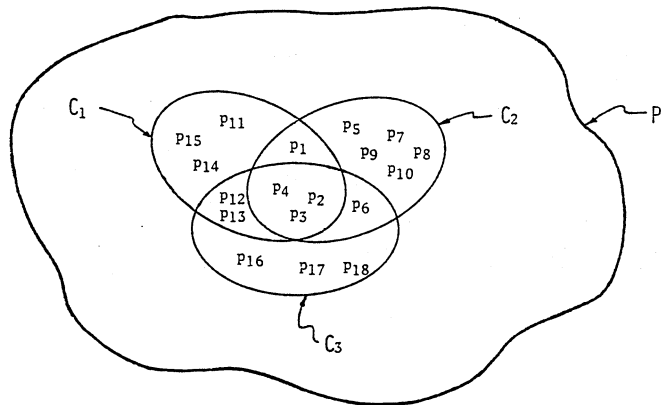
In the same fashion, as earlier derived, we may state the following:

MARKET (List of Products)      CONSENSUS LEVEL (1- $\alpha$ )

1.  $P_2, P_3, P_4$       100%
2.  $P_1, P_2, P_3, P_4$       66.7%  
OR  $P_2, P_3, P_4, P_6$   
OR  $P_2, P_3, P_4, P_{12}, P_{13}$
3.  $P_5, P_7, P_8, P_9, P_{10}$       33.3%  
OR  $P_{11}, P_{14}, P_{15}$   
OR  $P_{16}, P_{17}, P_{18}$

FIGURE 4

THE MARKET - THE THREE CONSUMER P-PRODUCT CASE

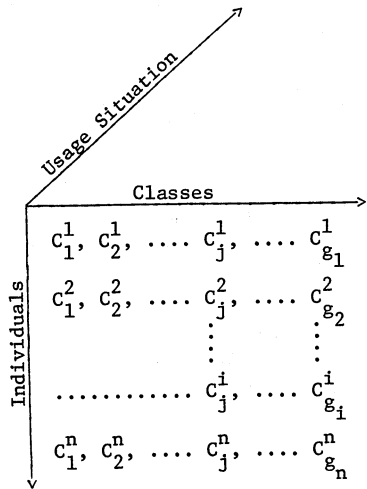


There are obviously other product combinations but the above combination serve to illustrate the basic concept of a delineated market and its associated confidence level. The decision maker is then in a position to define his relevant market with a predetermined level of confidence.

A Statistic for the General Case

The previous two simple examples serve to illustrate the procedure to be followed. In the general case, we wish to identify a group of products which are perceived as substitutes by (1- $\alpha$ ) of the consumers. For instance, the decision maker may decide that he wishes to delineate a set of products which are perceived as similar by at least 80% (i.e.,  $\alpha = 0.20$ ) of the population. If he were to draw a random sample of the population, the suggested method would yield objective and statistical conclusions for the choice of his market.

The statistic is derived from the information provided by each respondent. This information consists of a set of classes (groups of products) for each respondent for each usage situation. That is, the statistic is based on a measure of substitutability and is derived on the basis of each usage situation. These two dimensions have been suggested earlier by Massy (1975) and later emphasized by Day and Shocker (1976). The initial total list of products could be generated using Srivastava, Shocker and Day's (1977) suggested methodology. That is, each respondent presents his perceived markets (product substitutes) from a given set of P products. This could be represented by the following three dimensional data matrix:



where:  $i$ : identifies the  $i^{\text{th}}$  individual  
 $j$ : identifies the  $j^{\text{th}}$  class/group of products  
 $n$ : total number of individuals  
 $g_i$ : total number of classes defined by  $i^{\text{th}}$  individual  
 $C_j^i$ : the  $j^{\text{th}}$  class or set of products defined by the  $i^{\text{th}}$  respondent. As well:  
 $C_j^i = \{b_1^i, b_2^i, \dots, b_m^i\}$   
 where:  $b_m^i$ :  $m^{\text{th}}$  brand included by the  $i^{\text{th}}$  individual in the  $j^{\text{th}}$  group or class.

From the observed data we can define the following statistic:

$$P_k = \frac{\text{no. of individuals } \in \left\{ \begin{array}{l} \text{one of their } C_j^i \text{ contain the} \\ \text{same } k \text{ substitute brands} \end{array} \right\}}{n} \quad (3)$$

In the case where  $k$  (number of products in the supplied product set) is small, the calculation of all possible  $P_k$ 's becomes quite arduous. That is, if one wishes to see how  $k$  products out of  $P$  products could be combined to form various market definitions, then ( $P_k$ ) comparisons must be made before choosing that combination of  $k$  products which is most likely to occur.

For instance, if one had a total set of 40 products while considering delineating a market with 10 of these products, one would have to evaluate 847,660,528 (or  $P!/(k!(P-k))$ ) combinations before choosing the most likely combination, that is, that combination with the greatest  $P_{10}$ . In many cases, the exercise would probably be repeated for  $P_9, P_8, \dots, P_2$ . This would require the calculation of other likelihoods. It is obvious from this simple (and not abnormal) example that the computer time involved in calculating such probabilities would be prohibitive. Given these real constraints, a heuristic is needed which could shorten the time required to calculate each of these likelihood estimators. The following heuristic is seen as a proper approximation.

$$\hat{P}_k = \prod_{i,j \in k} P_{ij} \quad i \neq j \quad (4)$$

where:  $\hat{P}_k$ : estimated probability that a group of  $k$  products out of a given set of  $P$  products are perceived as substitutes

$P_{ij}$ : probability that products  $i$  and  $j$  are perceived as substitutes.

The great advantage with this approximation, besides that of being expeditious, is that we need only the initial probability matrix (i.e., the similarity measure for all pairs of products). A disadvantage is that it is an approximation and that it would tend to understate the actual  $P_k$ . On the other hand, it is only an estimator and the calculation of the "actual"  $P_k$  would then be obtained following the selection of the proper combination of products, which meets the  $1-\alpha$  predetermined level. The selection of the proper combination of products is that group of  $k$  products with the largest  $P_k$ . The actual selection is made by choosing the greatest  $\hat{P}_k$ . In order to verify for possible selection errors, given that we are using an estimator, one could calculate the actual  $P_k$  for those  $\hat{P}_k$  neighboring  $\text{Max}\{\hat{P}_k\}$  and then select the maximum value of  $P_k$ .

#### Procedure

Groups start forming in pairs of products. The degree of substitution between any two products is directly related to, and calculated from, the number of times any two products appear in the same class across all of the individuals. These are likelihood estimators for the proportion of people who see two products as being substitutes. These estimators are stored in a similarity matrix. Each element in this matrix is in fact an actual  $P_2$  for each pair of products.

The first group is formed by selecting the largest  $P_2$  element. The decision maker would have already set his confidence level at the  $\alpha$  level; we would then compare our maximum  $P_2$  to this  $1-\alpha$  level. If  $\text{Max}\{P_2\}$  equals or exceeds  $1-\alpha$ , we continue our search for additional members; if  $\text{Max}\{P_2\}$  is below the predetermined level, the procedure ends and we conclude that we have a homogeneous group of products at our predetermined confidence level.

If the procedure is allowed to continue, the two products with maximum  $P_2$  are chosen as a departure node for increasing the market size until we reach the desired cut-off point. Thus, the next product most likely to join the market will be that product which produces  $\text{Max}\{\hat{P}_3\}$ . For instance, if we assume that product number 5 will be chosen, among 40 products, to join products 11 and 23, then we have:

$$\hat{P}_3 = P(11, 5) \cdot P(23, 5) \cdot P(11, 23) \quad (5)$$

Generally, we have in the  $k = 3$  case:

$$\hat{P}_3 = P(i, j) \cdot P(i, m) \cdot P(j, m) \quad i \neq j; i \neq m; j \neq m \quad (6)$$

$$P_3^* = \text{Max}\{\hat{P}_3\} \quad (7)$$

where:  $P(i, j), P(i, m), P(j, m)$ : is the degree of perceived substitution between any two products (e.g.  $i$  and  $j$ ). That is, the proportion of people who perceive, for instance, products  $i$  and  $j$  as substitutes.

$\hat{P}_3$ : is the vector of all joint probabilities.

$P_3^*$ : is the maximum value in the  $\hat{P}_3$  vector. That is, that value related to the combination of 3 products most likely to be perceived as substitutes.

This procedure is extended to include whatever number of products the researcher requires in a market or until the cut-off point is reached, whichever occurs first. Thus, new members to the market are added while maintaining a predetermined level of market homogeneity.

Once a market is completely defined, the procedure

involves going back to the original similarity matrix and choosing the next pair of products which are most similar. From this point, we repeat the above procedure. This iterative process is carried out until all pairs have been exhausted or until the next highest substitution index is less than the critical level ( $1-\alpha$ ).

#### Conclusions and Implications

The preceding discussions have attempted to present a framework with which to examine the problem of determining the relevant market and a methodology with which this framework may be implemented. It is expected that this will have contributed to a reduction in the problem of identifying competitive market boundaries, and that the methodology presented will provide a defensible approach to defining a market. The true test of the proposed approach lies in its application - our following task.

The relevance of such research, for instance, to advertising is well illustrated by those anti-trust cases which have appeared before the courts. Advertising's role in the competitive process underlying our free market system has been a matter of controversy for many years. The importance of this subject for anti-trust regulations did not really emerge until the filing, by the FTC, of charges against the four leading ready-to-eat (RTE) cereal manufacturers (Kelloggs, General Mills, General Foods, Quaker Oats) under section 5 of the FTC act. It was claimed that their advertising practices result in high barriers to entry and is the main reason for the monopoly situation in this industry.

The case being made actually makes two basic and very important assumptions. First, it assumes that advertising can be a barrier to entry into the market. Second, implicitly assumed is a "defined" market. Depending upon how narrowly or broadly the market is being defined, the case for barriers to entry is more or less valid and thus the entire case actually lies in properly delineating the relevant market. Thus, it was the purpose of this paper to provide an objective means to operationalize this delineation task, while the implications of such research could be broadened to other areas of concern such as: evaluation of market share, product positioning, new product development, legal suits involving monopoly cases and the effective use of advertising.

#### References

- F.M. Bass, E.A. Pessemier, D.J. Tigert, "Complementary and Substitute Patterns of Purchasing and Use", Journal of Advertising Research, 9 (June, 1969), pp. 19-27.
- J.R. Bettman, "Decision-Net Models of Buyer Information Processing and Choice: Findings, Problems and Prospects", in Buyer/Consumer Information Processing, C.D. Hughes and M.L. Ray (eds.), Chapel Hill, North Carolina: U. of North Carolina Press, 1974, pp. 59-74.
- J.R. Bettman, "The Structure of Consumer Choice Processes", Journal of Marketing Research, 8 (Nov., 1971), pp. 465-471.
- J.C. Bourgeois, G.H. Haines Jr., M.S. Sommers, "Defining an Industry", presented at the joint national conference of the Operations Research Society of America (ORSA) and The Institute of Management Sciences (TIMS), Las Vegas, Nevada (Nov., 1975).
- Business Week, "Making the Switch to Something Else", Sept. 14, 1974, pp. 67-70.
- G.S. Day, A.D. Shocker, "Identifying Competitive Product-Market Boundaries: Strategic and Analytical Issues", Working Paper 156, Graduate School of Business, University of Pittsburgh (1976).
- J.N. Dean, "Product-Market Definition Under the Sherman and Clayton Acts", University of Pennsylvania Law Review, 110 (April, 1962), pp. 861-878.
- G.H. Haines Jr., "Process Models of Consumer Decision-Making", paper presented at the Association for Consumer Research Workshop in Information Processing, University of Chicago, 1972.
- G.H. Haines Jr., "Process Models of Consumer Decision-Making", in G.D. Hughes and M.L. Ray, Buyer/Consumer Information Processing, Chapel Hill, N.C.: University of North Carolina Press, 1974.
- J.A. Howard, J.N. Sheth, The Theory of Buyer Behavior, Wiley, 1969.
- D. Krech, R.S. Crutchfield, Theory and Problems of Social Psychology, McGraw-Hill, 1948.
- W.F. Massy, "Some Thoughts About the Definition of 'Relevant Market' in Antitrust Cases", Working Paper, Stanford University, 1975.
- W.T. Moran, "Why New Products Fail", Journal of Advertising Research (April, 1973).
- W.G. Nickels, Marketing Principles - A Broadened Concept of Marketing, Prentice-Hall, 1978.
- R.K. Srivastava, A.D. Shocker, G.S. Day, "An Exploratory Study of the Influences of Usage Situation on Perceptions of Product Markets", Working Paper No. 239, Graduate School of Business, University of Pittsburgh, 1977.
- V. Stefflre, New Products and New Enterprises: A Report on an Experiment in Applied Social Science, University of California, Irvine, California, 1971.
- P.O. Steiner, "Markets and Industries", in D.L. Sills (ed.), International Encyclopedia of the Social Sciences, New York: Free Press, 1968 (pp. 575-581).
- R.H. Strotz, "The Empirical Implications of a Utility Tree", Econometrica, 25 (April, 1957), pp. 269-280.
- R.H. Strotz, "The Utility Tree - A Correction and Further Appraisal", Econometrica, 27 (July, 1959), pp. 482-488.
- R.W. Wreth, "Determination of the Relevant Product Market", Ohio State Law Journal, 26 (Spring, 1965), pp. 241-293.

ATTRACTING POTENTIAL SWITCHERS TO MASS TRANSIT:  
MODE CHOICE AS A MULTI-ATTRIBUTE DECISION MODEL<sup>1</sup>

Linda L. Golden, John F. Betak, and Mark I. Alpert  
The University of Texas at Austin

Abstract

Policy makers and transportation planners frequently have to make decisions regarding the characteristics to incorporate into new or revised mass transit systems. To ensure consumer patronage, it is necessary to identify the trade-offs which are most likely to attract potential consumers. This paper uses conjoint analysis to identify such trade-offs and presents results and policy implications of using these procedures in one medium-sized American city.

Introduction

A frequent objective of transportation planners is to attract persons from low density modes (private automobile) to high density modes (mass transit). Given that the planner has some flexibility in determining the system attributes or characteristics, and given a desire to attract patrons of the system beyond those which might be considered a captive market, it becomes important to look to the potential consumer.

Due to limited resources, the planner and policy maker may not be able to incorporate all the characteristics into the system that the consumer finds necessary. However, just as the planner and policy maker must accept trade-offs, the consumer is also willing to accept trade-offs. And, with acceptable levels of certain attributes, although the consumer may not consider the system optimum, the consumer may patronize the system when he or she would not have otherwise. In a less than perfect world, it then becomes important to identify the trade-offs among system characteristics that are most likely to attract the consumer. With knowledge of these trade-offs, and knowledge of how the public views current characteristics of the automobile and mass transit, policy and planning decisions may be facilitated.

The focus of this paper is upon identifying trade-offs among levels of system attributes for potential users of mass transit and the resulting policy implications. The problems are couched in terms of a mode choice situation whereby the consumer has mode choice options and makes mode choice decisions on the basis of characteristics of alternative modes of transportation. In addition, this paper uses a methodology, conjoint analysis, which may be of considerable use to transportation planners. As Green and Srinivasan (1978) state, "Although conjoint analysis has been extensively applied mainly in the private sector, it has a large potential for public sector applications as well". While it is not new to stress the need to identify system attributes relevant to consumer mode choice decisions in an attempt to model and promote public transportation usage (Blattberg and Strivers 1970; Golob, Dobson and Sheth 1973; Hill and Von Cube 1963; Hill and Martin 1967; Mundy, Cravens and Woodruff 1974; Sheth 1975; Tybout, Hauser and Koppleman 1978), this new application of conjoint analysis may facilitate a further understanding of consumers' transportation decisions; hence, more efficient system design and potentially broadened consumer acceptance.

<sup>1</sup>This study was financed by the U.S. Department of Transportation under contract DOT-OS-30093 to the Council for Advanced Transportation Studies, University of Texas.

Mode Choice Problem

Mode choice is manifest by the behavior of individuals and/or groups of individuals. This is predominately a purposive, adaptive behavior. As such, mode choice behavior represents the outcome of a complex decision process which encompasses: (1) trade-offs between system characteristics and non-system characteristics, including user requirements and attributes; (2) past decisions with respect to mode choice, origins and destinations, life style, etc.; and (3) future decisions with respect to origins and destinations, life style, etc., i.e., goal-directed decisions.

It is the view of this study that mode choice decisions are one part of a large decision-making system in which each part of the system affects, and is affected by, the other decision components. A decision-making system may be conceptualized as having three main parts: "(1) a set of decision variables, (2) interrelationships among the decision variables, and (3) criterion functions depending on values of the decision variables" (Whinston, 1966). Given this orientation, this paper characterizes mode choice as a multiple criteria decision task. And, the problem of determining utilities in mode choice is considered to be one of determining part-worths of the mode's attributes at different levels.

Conceptualizing mode choice as a multiple attribute decision problem in which the individual chooses among alternatives which are described in terms of their attributes, it is possible to formalize the mode choice task as a multi-attribute utility model with the following essential features:

$$u(x) = c_1 u_1(x_1) + c_2 u_2(x_2) + \dots + c_n u_n(x_n)$$

where

$$x = (x_1, x_2, \dots, x_n).$$

A vector that specifies a consequence with  $n$  attributes, and  $x_i$  specifies the  $i$ th attribute.  
( $x$ ) = Set of possible consequences. The probability distribution of  $x$  depends on the decision maker's action choice.

$u(\cdot)$  = Utility function defined on ( $x$ ).

$U_i(\cdot)$  = Utility function defined on ( $x_i$ ).

$c_i$  = Scaling or weighting factor.

This is the so-called additive utility model which is discussed in some detail by Willie and Pessemier (1973). Using this model as normative framework for mode choice, it is argued that the individual will choose the transportation alternative that maximizes expected utility.

Given this specification of a model of mode choice decision-making, it now remains to determine a framework for obtaining utility values for a set of mode choice attributes. Conjoint analysis is a technique which decomposes the respondents' overall evaluations into separate, and compatible utility scales by which original global

judgments can be reconstructed. This contrasts with the compositional approach of expectancy-value models. Conjoint procedures require only rank-ordered output, yet yield interval scaled output. In the case of finite data, the scale is technically an ordered metric, but as the number of input values increases a unique representation at the interval scale level is approached (Green and Rao 1971).

Thus, conjoint measurement is concerned with the joint effect of two or more independent variables on the ordering of a dependent variable. For example, one's preference for various modes of transportation may depend upon the joint influence of such variables as cost, travel time, convenience, dependability, privacy and so on. Algorithms and applications of conjoint analysis have been discussed in the literature (Carroll 1969; Coombes, Dawes and Tversky 1970; Green and Carmone and Wind 1972; Green and Rao 1971; Johnson 1974; Krantz 1964; Kruskal 1965; Luce and Tukey 1964; Srinivasan and Shocker 1973; Tversky 1967; Westwood, Lunn and Beazley 1974; Young 1969), and the purpose of this paper is not to discuss the mathematics of the methodology, but to investigate its usefulness in a modal choice situation. Green and Wind (1975) provide a discussion of the technique, for those unfamiliar with it, and Green and Srinivasan (1978) provide an excellent paper on issues and outlook for consumer behavior.

#### Methodological Development

In order to identify those persons who were most likely to switch from the automobile to mass transit and the attributes to be used in the trade-offs analysis, a pilot study was conducted. The pilot study consisted of 252 usable personal interviews obtained from a stratified by census-tract (quotas proportional to population) area random sample of households in a medium-sized southwestern city. The purpose of this study was to identify determinant attributes (Myers and Alpert, 1968) for mass transit and demographic characteristics of potential switchers to mass transit.

#### Selection of Attributes

The results of the pilot study indicated that 11 attributes of the 27 tested were determinant (both perceived as important and varying across modes, in this case the bus and car). These were: dependability, low energy use per passenger, economy, low pollution per passenger, convenience, flexibility, freedom from repairs, freedom from accidents, no parking problems, brief travel time, and safe from dangerous people.

On the basis of the above results and extensive literature reviews (Betak and Betak 1969; Davies and Alpert 1975; Alpert, Golden, Betak and Story 1977), nine attributes of transportation modes were selected for trade-off evaluation: cost per mile, fuel use per passenger, transportation available--hours per day, travel time is--minutes, possibility of encountering dangerous people, level of comfort, opportunity to socialize, and transportation available--days per week. With the exception of level of comfort and opportunity to socialize, the attributes used for the trade-off analysis were determinant. These non-determinant attributes were reviewed in the literature and may be operationalized by policy makers to improve the system.

To operationalize some of the determinant attributes it was useful to redefine them in terms which could be related to observable phenomena. For example, the attributes of dependability, flexibility, and convenience were operationally defined in this study to mean "transportation available \_\_\_\_\_ hours per day" and "transportation available \_\_\_\_\_ days per week." To

operationally define the attributes of economy and energy, "cost per mile" and "fuel use per passenger" were utilized. To operationalize the attribute of brief travel time, "total travel time is \_\_\_\_\_ minutes" was used. No attempt was made to provide operational definitions of comfort, safety from dangerous people or socializing.

Each attribute was conceptualized as a three-level variable. In the case of "cost per mile," the levels were defined as being present cost, 15¢ less than present cost, and 15¢ more than present cost. To assist the respondents in calculating their present cost, estimates of typical current operating costs of an automobile or a bus ride were provided in the introduction of the questionnaire. The attribute of "level of pollution per passenger" was defined as low, medium and high. The levels of "transportation available \_\_\_\_\_ days per week" were defined as Monday through Friday or five, Monday through Saturday or six, and Monday through Sunday or seven. The levels of "transportation available hours per day" were defined as twelve, eighteen, and twenty-four. "Total travel time is \_\_\_\_\_ minutes" was defined as fifteen, thirty and sixty minutes (Harman 1974; Redding 1970). "The possibility of encountering dangerous people" was defined as never, sometimes, and often. The attribute "level of comfort" was defined as having three levels of low, medium, and high. The attribute of "opportunity to socialize" was defined as having three levels of never, sometimes, and often. "Fuel use per passenger" had three levels of low, medium, and high.

#### Sample of Potential Switchers

The pilot study identified potential switchers as those who do not now ride the bus, but said they would definitely ride if it were improved. Ten percent of the random sample answered definitely yes to this question, and thus, were identified as potential switchers. These people tend to be relatively younger, have smaller households, are most likely to be full or part-time students, are more likely to work or shop downtown, and tend to be higher educated than non-potential switchers. Potential switchers did not differ from the aggregate sample by income, occupation, number of cars or race.

On the basis of this information, eleven areas of the city were identified from census tract data as having a high proportion of potential switchers who live within one-quarter mile of mass transit routes. An enumeration of households in these areas was obtained from Cole's Directory. Computer-generated random numbers were used to identify every nth person to be included in the sample frame. Three hundred and seventy-five potential respondents were mailed personal letters asking for their participation in the study which was followed by a phone call from the interviewer establishing a date, time and place for the interview. Of the 201 people contacted by phone, 60 agreed to participate, resulting in 48 usable personal interviews.

#### Final Instrumentation

The nine attributes and their levels were developed as pairwise matrices. The thirty-six matrices, each with three levels by three levels trade-offs for each pair of attributes, requires a respondent to make 324 rankings (36 times 9 rankings per matrix). To facilitate this task, illustrative graphics were utilized throughout the matrices on the final questionnaire to characterize the attributes and their levels. In a pilot study, it had been shown that these graphics increased respondents' speed of completing the rankings without altering the relative utilities obtained for levels of the attributes rated (Alpert, Golden, Betak, Story and Davies 1977). Three separate pre-tests were used to refine the instrument and test the usefulness of graphic and



verbal presentation of attribute levels.

The final instrument provided a set of instructions, "warm-up" matrices, trade-off matrices, attribute level perceptions for bus and private car, demographics and questions concerning current ridership patterns. Respondents were asked to respond in the frame of reference of trips to work or school. (Trade-off matrices elicited information about the importance of attribute levels for a transportation made in general, not for a specific mode). Interview time averaged forty-five minutes.

A parallel study (using the same sampling procedures) was undertaken, forming the trade-offs into an alternative method for collecting data, in which unique combinations of attributes and levels are placed on separate cards, each representing a possible combination that might be available. The potential advantage of the "card-sort" approach is that it illustrates the "Gestalt" and has respondents make choices involving wholes rather than pairs of attributes and levels (Green and Wind 1975). Once the data are obtained, it is possible to analyze card-sort data in the same manner as matrix data, since an orthogonal array is used to generate equivalent combinations of levels of attributes to be trade-offs against other combinations, yielding similar choice problems for levels of attributes (Johnson 1973 and 1974). Using an orthogonal array substantially reduces the number of combinations that are needed, in this case, from 3<sup>9</sup>, or 19,683 combinations to 27 combinations that are minimally necessary to secure trade-offs equivalent to those obtained in the matrix approach, also assuming no interactions among attribute levels.

However, even with this limited amount of cards, the 50 respondents involved in the "card sort" procedure were faced with an information-overload situation. Measures of the ability of the algorithm to capture the utilities sufficient to reproduce accurately the rankings given showed the card-sort procedure to be substantially less reliable than the matrix one, FOR THIS NUMBER OF ATTRIBUTES AND LEVELS (Alpert, Golden, Betak, Story and Davies 1977). Accordingly, we present results from the 60 matrix method respondents whose demographics correspond closely to those of previously identified potential switchers.

#### Analysis and Discussion of Results

Data analysis evolved through several stages. The first stage investigated the validity of the matrix data. The second stage involved determination of utility curves and linear equations for the trade-off data, and the third state of analysis provided information from which policy statements could be made concerning attribute trade-offs for the car and the bus.

#### Validity Measures

Two types of analysis were used to evaluate the validity of the utility model mode choice proposed. The first analysis used  $\theta$  (theta) to evaluate the goodness of fit of the data by examining the relationship between the input rank order of the data and the obtained rank order as derived from the trade-off algorithm. The algorithm used is the non-metric regression analysis developed by Johnson (1973). The lack of fit measure is  $\theta$ , where:

$$\sum_{i < j} \delta_{ij,kl} (d_{ij}^2 - d_{kl}^2)^2$$

$i < j$

$$\theta^2 = \frac{\sum_{\substack{k < l \\ i < j \\ k < l \\ (i,j) \neq (k,l)}} (d_{ij}^2 - d_{kl}^2)^2}{\sum_{\substack{k < l \\ i < j \\ k < l \\ (i,j) \neq (k,l)}} (d_{ij}^2 - d_{kl}^2)^2}$$

and

$$\delta_{ij,kl} = \begin{cases} 1 & \text{if sign}(d_{ij} - d_{kl}) \neq \text{sign}(r_{ij} - r_{kl}) \\ 0 & \text{otherwise} \end{cases}$$

For pairwise, two attribute trade-offs,  $\theta$  will be zero if the  $d_{ij}$  have the desired rank order, and unity if their order is perfectly reversed (Johnson 1973).

Table 1 presents  $\theta$  values for selected control groups. Eight categories of controls are used. In the first category  $\theta$ 's from all respondents were analyzed. In the second category respondents were grouped into three classes on the basis of a post-interview evaluation of their seriousness and level of effort in completing the instrument. The post-interview evaluation was done by a non-interview team using the remarks of the interviewers written on each instrument. In this second category, the first three quality levels of the respondents were grouped together and their data submitted for analysis. The remaining controls were for sex, age, and satisfaction levels, all using the first three quality levels of respondents. The satisfaction category is limited to those respondents who are very satisfied with their present mode of transportation.

As can be seen from Table 1, the  $\theta$  values ranged from .327 to .142. The  $\theta$ 's for the respondents are relatively low, indicating that the derived weights for the attributes are reasonably consistent with the input rank order data. In short, it is possible to interpret the rank ordering of the attributes of the respondents with some degree of surety that these weights are a meaningful representation of the part-worths of the attributes investigated.

TABLE 1  
COMPARISON OF  $\theta$ 'S FOR SELECTED GROUPS

	N	$\theta$
All respondents	60	.27709
Quality 1, 2, 3	48	.28227
Males, Quality 1, 2, 3	34	.28769
Females, Quality 1, 2, 3	14	.24531
Age 18-29, Quality 1, 2, 3	7	.14225
Age 30-44, Quality 1, 2, 3	17	.22119
Age 45+, Quality 1, 2, 3	24	.32703
Very satisfied, Quality 1, 2, 3	48	.26399

To consider further the issue of the validity of the results obtained, the consistency of the rank order of the attributes obtained was compared with the results of previous research. This issue was considered by comparing the range of weights obtained for each attribute, and the rank order of the attributes for all of the respondents. The difference between the weights (utilities) for the high and low levels of an attribute indicate how sensitive that attribute is to level

changes, hence, how much influence it may have in determining choice (see Table 2). A large range indicates that variation in the amount of the attribute available in a mode will significantly affect the utility of that mode in a choice situation. The attribute ranges were: Possibility of Encountering Dangerous People, 1.365; Energy Use, 1.234; Pollution, 1.228; Total Travel Time, 1.210; Cost, 1.066; Available Days/Week, .945; Available Hours/Days, .848; Level of Comfort, .670; Opportunity to Socialize, .506.

The data indicate that the rank order of the attributes derived from the trade-off questionnaire were generally consistent with other research (Betak 1969; Davies and Alpert 1975; Alpert, Golden, Betak, Story, Davies 1977). Some differences did occur and were expected given the nature of the task confronting the respondents, the sample and the particular set of variables used.

It is clear from the thetas and consistency with other research that meaningful interpretations of the trade-offs may be made. Subsequent discussion focuses on these analyses, and are restricted to the 48 respondents contained in the group of quality 1, 2, 3 interviews.

#### Attribute Level Utilities (Part-Worths)

The first form of analysis was a determination of the sample's utilities for each attribute in each pairwise trade-off. Due to the volume of data, Figure 1 illustrates a matrix presented to respondents and this form of analysis. In this figure, the sample's utilities for the attribute levels are indicated by the decimal values at the right and bottom of each matrix. The algorithm, as previously described, computes the joint additive utility for each attribute and level pair, and is indicated in the top part of the respective cells in the matrix. The utility of each combination is the sum of the two level's utilities in each matrix.

FIGURE 1

#### SAMPLE'S UTILITIES: COST VERSUS POLLUTION

COST PER MILE:

Low	1.16592 (1)	.88100 (2)	.51005 (4)	.79309
Medium	.57759 (3)	.29267 (5)	.07828 (6)	.20476
High	-.18076 (7)	-.46568 (8)	-.83663 (9)	.55359
	.37283	.08791	-.28304	

Table 2 provides a summary of the average utility or part-worth, given by the sample for each specified attribute level. The average was calculated as the arithmetic mean at all derived weights (as in Figure 1) for each level of an attribute, as determined through all possible trade-offs with all other attributes and levels.

As noted above, the rank order of the attribute determinances show an essentially similar pattern to those studies that have included similar variables. Time, convenience, and psychological safety variables tend to dominate over level of comfort and time of day availability, given current levels of these attributes on public transportation systems (vs. cars) as represented in the trade-off study. Caution should be taken to realize that these utility ranges are sensitive to the selected levels shown in Table 2; variations in ranges chosen might produce different rankings of relative determinance of attributes. Care was taken to select

levels which typified the relevant range for current transportation modes, but the utility ranges are nevertheless unique to the trade-offs here.

TABLE 2

#### SAMPLE'S UTILITY VALUES (PART-WORTHS)

Attribute	Level-Specified	Average Utility Value	Utility Range
Cost per mile	-15¢	.543	
	your present cost	.145	1.066
	+15¢	-.523	
Energy/use pssgr.	low	.679	
	medium	.166	1.234
	high	-.554	
Pollution/pssgr.	low	.698	
	medium	.179	1.228
	high	-.529	
Days/wk. service available	5	-.444	
	6	.143	.945
	7	.501	
Hrs./day service available	12	-.396	
	18	.199	.848
	24	.451	
Total travel time in minutes	15	.632	
	30	.186	1.210
	60	-.579	
Possibility of Encountering Dangerous people	never	.781	
	sometimes	.196	1.365
	often	-.585	
Level of comfort	low	-.289	
	medium	.267	.670
	high	.381	
Opportunity to socialize	never	-.155	
	sometimes	.351	.506
	often	.152	

Another general observation from Table 2 is that with one exception the part-worths are all relatively linear representations of increasing or decreasing utility, in intuitively appealing directions. For example, the utility of cost per mile decreased as cost increased, and the same was found for energy use per passenger, travel time in minutes, possibility of encountering dangerous people, and so forth. The reverse results were shown for service availability, in that utility increased as did availability. The linear equations which might prove useful for interpolating within these ranges, and making modal combinations with equivalent utilities yet differing costs to deliver, are described elsewhere (Alpert, Golden, Betak, Story, Davies, 1977).

The utilities for opportunity to socialize showed a relatively more interesting non-linearity, in the sense that utility peaked at the middle level, implying a "golden mean" for socializing (perhaps too much involves perceptions of crowding or obtrusive conversations). This non-linear relationship with modal satisfaction would be considerably more difficult to discern with the more conventional multiple regression approaches to utility estimation (Banks, 1950; Hughes, 1966).

#### Private Automobile versus Public Transportation

The final stage in the analyses concerned using the calculated utilities to assess how respondents viewed private automobile and public transportation at the time of the interview. Viewed in total, Table 3 has a great deal of information that can aid in policy decisions. Columns 1 and 2 provide perceived utilities

for the respondents' perceptions of the images of public and private transportation features. Assuming rationality, there is no reason to expect public transportation to be chosen, since the private auto clearly has the highest overall utility, given the present level of system attributes. Thus, the remainder of Table 2 provides guides to policy makers with respect to focal points for making changes in the modes of transportation to obtain increased patronage for one mode or another (typically trying to increase public transportation patronage).

The third column of Table 3 shows the relative advantage or disadvantage, per attribute, of public transportation's image. Five attributes of the nine show negative gaps for public transportation. Column 4 shows that both modes are quite distant from the maximum utility levels obtainable, across all attributes, given the perceived (and actual) characteristics.

One way for public transit to gain on private autos would be to improve toward the maximum utility level for each attribute. Column 5 indicates the potential gains that could be obtained, utilizing a positive or "carrot" (vs. "stick") strategy of improving the perceived image of public transportation to the maximum utility level. This column should be balanced against the relative economic costs of moving perceptions from their current point to or toward the maximum utility level. Rather than lowering revenues by instituting lower fares, for example, it might be less costly, and more effective to decrease the perceived probability of encountering dangerous people on public transportation.

counter to their expectations), and so forth. Several of these means are relatively inexpensive, yet may succeed in raising the perceived utility of this feature for enough of the .781 potential gain that the current overall disadvantage is largely offset. Although dial-a-ride would be considerably more expensive than lighting and promotional activities, this would also serve to increase perceived comfort utility, as well as that for total travel time, both of which Table 2 shows have high potential gain. Further research is needed to measure the dollar costs of moving perceptions along these attribute levels, and the incremental ridership levels of closing the gap (or surpassing) private transportation's utility for commuting.

Columns 6 and 7 suggest alternative policy and forecasting scenarios for public transportation's relative utility versus private autos. If, for example, a tax were levied on auto use sufficient to increase the price per mile to the lowest level of utility (column 6), the relative utility of buses would increase by .540, which is the difference between auto's current perceived utility (0.17) and the minimum utility in the scale range (-.523). If achieving positive gains may be viewed as the "carrot" approach, then imposing penalties on auto usage may be viewed as the "stick". Many planners have suggested more emphasis on the latter. While planners may wish to incorporate mostly positive improvements in the public transportation modes, to the extent that increased costs are enacted or may be forecasted (due to the increasing fuel costs of private cars, for example), the conjoint utility data presented here may prove useful in evaluating the quantitative

TABLE 3

COMPARISONS BETWEEN WEIGHTED UTILITIES FOR PUBLIC VS. PRIVATE TRANSPORTATION FOR COMMUTING TRIPS, BY ATTRIBUTES: POTENTIAL GAINS FOR PUBLIC TRANSIT VIA POSITIVE AND NEGATIVE STRATEGIES

Attribute	Weighted Utility Mass Transit	Weighted Utility Auto	Mass Minus Auto (gap)	Maximum Utility	Maximum Positive Gain	Minimum Utility	Maximum Negative Gain
Cost	.346	.017	.329	.543	.197	-.523	.540
Fuel Use/Pssgr.	.345	-.132	.477	.679	.334	-.554	.422
Pollution/Pssgr.	.304	.186	.118	.698	.394	-.530	.716
Days/Week	.335	.485	-.150	.501	.166	-.444	.929
Hrs./ Days	.019	.436	-.417	.451	.432	-.396	.832
Total Travel Time	-.171	.315	-.486	.632	.803	-.579	.894
Possibility of Encountering Dangerous People	.115	.252	-.137	.781	.666	-.585	.832
Comfort	.089	.329	-.240	.381	.292	-.289	.618
Opportunity to Socialize	.289	.265	.024	.351	.062	-.155	.420
	1.671	2.153	-.482	5.017	3.346	-4.055	6.203

<sup>a</sup>Gain in total utility for public transportation if mean perception were raised to highest utility level for particular attribute(s).

<sup>b</sup>Relative gain in total utility for public transportation if mean perception of auto were lowered to lowest utility level for particular attribute(s).

As this variable is in part a proxy for social class perceptions of riders (mainly by non-riders), part of the non-rider's fears can be countered by scheduling routes through "safe" neighborhoods. Additional strategies might involve providing between lighting at bus-stops, instituting dial-a-ride (essentially door-to-door) service, placing security personnel on buses, featuring persons in bus advertisements who are similar to the target market of potential switchers (and somewhat

impact of projected changes in either mode's features. Increased traffic congestion in inner cities, coupled with bus lanes and express service, may be forecasted as leading to net utility gains for public transportation's travel time dimension, where it currently has its greatest perceived disadvantage. However, persons attempting to manipulate such relative travel time by bus lanes should be aware of the potential factors that partially led to the abortive bus lane preferential treatment experience in Los Angeles during 1976 (Tischer and Shea 1978). Taking into account the political and environmental factors, column 7 can be useful in indicating alternative, albeit negative, approaches to influencing (or forecasting) improvements in the relative utility of public vs. private transportation.

Clearly, there are several combinations of policies or

environmental changes that may occur to change modal perceptions and/or perceived utility values for particular levels of attributes. Given relatively stable utilities for various modal attribute levels, Table 3 can prove useful in guiding policies for altering the relative advantage in favor of public transportation. In addition, it is recommended that repeated studies be done to check on changes in the perceived utilities of levels over time and changing conditions and social values. If we can assume that respondents would generally behave rationally and choose the most preferred alternative, and that public transportation were to gain in perceived utility for sizable numbers of respondents, significant shifts might be possible. If the appropriate promotional techniques were utilized to apprise the respondents of the alteration in the transportation system, and accounting for a certain amount of lag or inertia in choice behavior, an increasing utilization of public transportation would be expected for respondents having characteristics in common with the sample studied here.

#### Conclusion

Viewing modal choice decisions as part of a large decision-making system in which each part of the system affects and is affected by the other decision components a multi-attribute utility model can be used to describe the consumer's mode choice decision. Applying conjoint analysis to this framework, the researcher can identify utilities and trade-offs among system attributes for both mass transit and private transportation. This yields information which can assist the marketer and planner of public transportation in determining changes in the system most likely to increase ridership.

This study indicated that policies directed toward improving total travel time, service availability in hours per day and days per week, safety from dangerous people, and comfort would be most likely to improve the overall utility of public transportation. Further, ridership would be likely to increase if scheduling and headways were arranged to provide total travel time of 15 minutes, 24 hours per day, 6 days per week and at a cost of 15 cents more per mile than current cost. The market segment is willing to pay more for an improved combination of attributes designed to meet their needs.

If policies were directed to improving the utility for these attributes, or others with higher utility gains relative to costs of changes, and assuming that the respondents behaved rationally and chose the most preferred alternative, and that the characteristics of the automobile were not altered, then public transportation would have a perceived total utility higher than a private automobile. Alternatively, we have shown how changes imposed on auto characteristics by policy and/or environmental conditions may similarly affect this relative utility. Assuming effective promotional techniques, a campaign to inform individuals like those in the sample about the improvements in the public transportation system should yield an increasing utilization of public transportation.

Limitations on the type of analysis performed in this study are in assessing whether the policy options having the greatest potential for altering choice behavior were feasible politically or economically. Furthermore, these types of analyses cannot indicate which combinations of the changes in transportation attributes would yield the most cost-effective option. Also, no information was obtained for attributes outside the range presented in this study. Thus, analyses beyond the scope of this research must be undertaken to fully utilize the results reported herein.

From the preceding results, several suggestions for future research appear germane. First, incremental changes in the attributes having the greatest potential for altering utilities should be implemented and monitored. Second, analytical models for evaluating the political and economic viability of alternative attribute combinations for transportation systems need to be developed, possibly linear programming models. Finally, in conjunction with the recommended study of incremental changes, further development should be undertaken of more parsimonious instrumentation for eliciting trade-off data from potential users of transportation services, minimizing the respondents' time investment, and reducing the computational costs of analyzing trade-off data.

In conclusion, the trade-off analyses developed in this study provide indications of the areas where policy may be most effective in increasing the relative utility of public transportation services. These findings provide at least a first handle on some of the policy levers that may be available to decision makers confronted with choosing alternative strategies for the provision of public transportation in their communities.

#### References

- Alpert, M.I., Golden, L.L., Betak, J.F., Story, J., and Davies, C.S. (1977), An Evaluation of Promotional Tactics and Utility Measurement Methods for Public Transportation Systems, Research Report 39, Council for Advanced Transportation Studies, The University of Texas at Austin.
- Banks, S. (1950), "The Relationship Between Preference and Purchase of Brands," Journal of Marketing, 15, 145-157.
- Betak, J., and Betak, C. (1969), Urban Modal Choice: A Critical Review of the Role of Behavioral Decision Variables, Research Report, The Transportation Center, Northwestern University.
- Blattberg, R.C., and Stivers, R. (1970), "A Statistical Evaluation of Transit Promotion," Journal of Marketing Research, 7, 293-299.
- Carroll, J.D. (1969), "Categorical Conjoint Measurement," Ann Arbor, MI: Meeting of Mathematical Psychology.
- Coombs, C.H., Dawes, R.M., Tversky, A. (1970), Mathematical Psychology, An Elementary Introduction. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.
- Davies, S., and Alpert, M.I., (1975), Modal Choice and The Value of Passenger Travel Time Literature: A Selective Bibliography, Research Report 22, Council for Advanced Transportation Studies, The University of Texas at Austin.
- Golob, T., Dobson, R., and Sheth, J.N. (1973), "Perceived Attribute Importance in Public and Private Transportation," American Institute for Decision Sciences Proceedings, 5, 7-10.
- Green, D.E., Carmone, F.J., and Wind, Y. (1972), "Subjective Evaluation Models and Conjoint Measurement," Behavioral Science, 17, 288-299.
- \_\_\_\_\_, and Rao, V.R. (1971), "Conjoint Measurement for Quantifying Judgmental Data," Journal of Marketing Research, 8, 355-363.

- \_\_\_\_\_, and Srinivasan, V. (1978), "Conjoint Analysis in Consumer Research: Issues and Outlook," Unpublished research paper.
- \_\_\_\_\_, and Wind, Y. (1975), "New Way to Measure Consumers' Judgements" Harvard Business Review, 53, 107-117.
- Harman, E.J. (1975), "A Behavioral Analysis of the Concepts Used in Housing Choice" Ph.D. Thesis, Department of Geography, McMaster University.
- Hill, D.M., and Von Cube, H.G. (1963), "Development of a Model for Forecasting Travel Mode Choice in Urban Areas," Highway Research Record, 38, 78-96.
- Hille, S.J., and Martin, T.K. (1967), "Consumer Preferences in Transportation," Highway Research Board Record, 42, 6-43.
- Hughes, G.D. (1966), "Developing Marketing Strategy Through Multiple Regression," Journal of Marketing Research, 3, 412-415.
- Johnson, R.M. (1973), "Pairwise Nonmetric Multidimensional Scaling," Psychometrika, 38, 11-18.
- \_\_\_\_\_, (1974), "Trade-off Analysis of Consumer Values," Journals of Marketing Research, 11, 121-127.
- Krantz, D.H. (1964), "Conjoint Measurement: The Luce-Tukey Axiomatization and Some Extensions," Journal of Mathematical Psychology, 78, 248-277.
- Kruskal, J.B. (1965), "Analysis of Factorial Experiments by Estimating Monotone Transformations of the Data," Journals of the Royal Statistical Society, Series B, 27, 251-263.
- Luce, R.D. and Tukey, J.W. (1964), "Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement," Journal of Mathematical Psychology, 78, 1-27.
- Mundy, R.A., Cravens, D.W., and Woodruff, R.B. (1974), "Potential for Marketing Management Applications in Public Transportation Planning," American Marketing Association Proceedings, 274-279.
- Myers, J.H., and Alpert, M.I. (1968), "Determinant Buying Attitudes: Meaning and Measurement," Journal of Marketing, 32, 13-20.
- Redding, M.J. (1970), "The Quality of Residential Environments: Preferences for Accessibility to Residential Opportunities," Unpublished Ph.D. dissertation, Department of Civil Engineering, Northwestern University, Evanston, Illinois.
- Sheth, J.N. (1975), "A Psychological Model of Travel Mode Selection," Association for Consumer Research Proceedings, 2, 425-430.
- Srinivasan, V., and Shocker, A.D. (1973), "Estimating the Weights for Multiple Attributes in a Composite Criterion Using Pairwise Judgments," Psychometrika, 38, 473-493.
- Tischer, M.L., and Shea, C.G. (1978), "Attitude-Behavior Changes in a Before-After Mode Choice Situation," Association for Consumer Research Proceedings, 5, 456-464.
- Tversky, A. (1967), "A General Theory of Polynomial Conjoint Measurement," Journal of Mathematical Psychology, 81, 1-20.
- Tybout, A.M., Hauser, J.R., and Koppelman, F.S. (1978), "Consumer Oriented Transportation Planning: An Integrated Methodology for Modeling Consumer Perceptions, Preference and Behavior," Association for Consumer Research Proceedings, 5, 426-434.
- Westwood, R., Lunn, T., and Beazley, D. (1974), "The Trade-off Model and Its Extensions," Journal of the Market Research Society, 16, 227-241.
- Whinston, A. (1966), "Theoretical and Computational Problems in Organizational Decision-Making," in J.R. Lawrence (ed.) Operational Research and Social Sciences. London, New York: Tavistock Publishers, 191-207.
- Wilkie, W.L., and Pessemier, E.A. (1973), "Issues in Marketing's Use of Multi-Attribute Attitude Models," Journal of Marketing Research, 10, 428-441.
- Young, F.W. (1969), "Polynomial Conjoint Analysis of Similarities: Definitions for a Special Algorithm," Research Paper No. 76, Chapel Hill, N.C.: University of North Carolina, Psychometric Laboratory.

Ivan Ross, University of Minnesota

## Introduction

The three papers composing this set share some things. First, each has a public policy implication (but, doesn't everything these days?). Second, each proposes or implies methodological procedures for estimating the effect of certain market actions (e.g., changes in attributes, repositioning of products) on consumer responses, perceptions, or behavior (Burger and Venkatesh and Golden, *et. al.*) or, vice-versa (Bourgeois). Indeed, the CB radio and transportation papers are very closely related, and were that it were possible for the authors of these two papers to have spent a day in seclusion before this Conference, we might have heard today one instead of two papers, and moreover, some of the confusion this discussant retains concerning the relative merits or demerits of alternative procedures for estimating attribute utilities (trade-off analysis, factor and discriminant analysis, etc.) might have been ameliorated as a result of such a coalition.

Third, each of the papers, in its own way, makes some advance in either methodological or conceptual thinking and application by extending our knowledge in three areas where, by this time, there is an accumulating "history" of research.

A Measure for Market Delineation  
Jacques C. Bourgeois

## What is a Market?

There is no question but that procedures for determining or estimating the extent to which multiple "offerings" define a set within which high substitutability obtains (high cross-elasticities of demand) are useful. Moreover, as Bourgeois notes, there are noteworthy differences in approach to this question by various investigators. The troublesome thing about the conceptual positioning of this article, however, is that Bourgeois' proposed procedure for "defining" a market appears to be nothing more than a perceptual measure of substitutability which is offered as a "substitute" for a behavioral measure of substitutability. It is unclear how this advances the state of our knowledge or of our methodological repertoire.

Moreover, the author does not directly confront the question of the adequacy of a measure of substitutability as a measure of "market." Are not there many meanings of "market" which have relevance which do not directly depend upon any conception of cross-elastic demand? For example, in trademark infringement matters, is it not the case that confusion as to source or origin between two products has a "market definition" meaning, whether or not consumers perceive the products as substitutes?

Naturally, the author is free to employ any definition of "market" he wishes to demonstrate the value of the heuristic he offers to aid in its delineation, but some would consider it useful to read a considered evaluation of how this definition of market is most usefully applied or is more pertinent in terms of its applications.

## Individual Versus Aggregate Analysis Levels

The terms "market" and "market segment" are being used in Figures 3 and 4 in a way that may lead to some con-

fusion. This discussant reads the notation as defining the "market" as consisting of "all and only those" products seen in common by consumers as satisfying the same "need", and as defining "market segments" as those products seen by at least one, but not all consumers, as satisfying the same "need." But then, apparently, Figures 3 and 4 do not contain any products which are not part of either a market or a market segment (i.e., none is outside of a "C set"). Thus, when Bourgeois states that (with reference to Figure 3), "the present market is therefore defined as the set of those four available products...", he is restricting our attention to the "market" and not to the "market segments," and seems to be asserting that only these four products compose a "market" for these two consumers. This discussant sees fifteen products which compose "a market" for these two consumers, and he is 100% certain of that. That four of them "compose a market" for the same two consumers is undoubtedly of marketing interest but is not an entirely satisfactory basis for asserting we are "less certain" that the other eleven products are not "in" the market for "that" need. There may be a semantic problem here resulting from moving between "individual" (consumer) and "aggregate" (market) levels of analysis.

## The Value of the Proposed Market Delineation Heuristic

The approximation heuristic (equation 4) estimating the probability that a particular set of products are perceived as "substitutes" seems an efficient procedure for identifying "similar" product groupings. It may be of interest to inquire, however, whether or not consumers understand (perceive) the concept of "market" and the product set that composes it (them) to be determined simply and exclusively by whether or not they place the same two or more products in the same class/usage situation cell. And how does the calculation of aggregate probability values handle obvious problems of segmentation with respect to product groupings/sets and interactions of such groupings with usage situations? Can the heuristic be used for investigating the existence of segments and the marketing implications of repositioning products (given some measurement on attributes which form the basis for a consumer's placement of a product in a class/usage situation)?

It would appear that the heuristic can be quite useful in various managerial applications, but it would be helpful if the author would provide examples of applications which would shed light on issues such as... what levels of confidence for "homogeneity" are relevant/required for what kinds of decisions... what if the same product belongs to several segments... how do we bridge between probability estimates and product/individual attributes which cause product sets to occur in the first place?

Attracting Potential Switchers to  
Mass Transit: Mode Choice as a  
Multi-attribute Decision Model

Linda L. Golden, John F. Betak, and Mark I. Alpert

### Individual Versus Aggregate Analysis Levels

Although the paper very usefully extends our methodological "tool kit" in looking at choice behavior in the mass transit situation, some will have a concern about the loss of individual level data as a result of calculating attribute utilities across subjects. Ultimately, if we want to predict the number of persons who might be moved from personal auto to mass transit via the change or addition of some determinant attribute or combination thereof, don't we have to retain data at the individual level and in some way simulate choice changes per each individual via manipulation (real or hypothetical) of attributes/levels of attributes?

### Definitional Problems

Several issues deserve clarification. How did the authors begin with 27 attributes from which 11 were identified as (presumably), "determinant", and end up with using only nine in the trade-off analysis, and of these nine, two were "non-determinant"? Some of the "operational definitions" seem equivocal. For example, the number of hours per day and number of days per week that a transportation mode is available may well be a part of attributes such as "dependability", "flexibility", and "convenience", but would not intuitively seem to exhaust the definitions of these attributes. Does not "dependable" suggest more than the number of hours/days available? How "flexible" or "convenient" is a transportation mode if it is available eighteen hours a day but not at all between those hours in the morning or afternoon when a particular rider may require its use?

These issues would appear rather crucial since the information with which we are presented in Tables 2 and 3 presumably is sufficient to lead the authors (and readers) to decisions as to which attributes "need" changing most in order to increase the probabilities of mass transit patronage. Therefore, if there are problems in the way in which determinant attributes have been selected and defined, this information may be misleading.

Further, since the outcome of the conjoint analysis is so sensitive to the particular levels chosen to represent each attribute, the rationale for this procedure seems to deserve greater attention. How does one determine if the "levels" chosen to illustrate an attribute are reasonable, realistic, and "equivalent", so to speak, across attributes.

Not only does this issue receive scant attention in this research paper, it seems a common deficiency in most of the conjoint literature. To argue as do the authors that "care was taken to select levels which typified the relevant range for current transportation modes..." may not be a satisfactory assurance for many researchers. Consider, for example, the possible difference in part-worth utility attaching to a range of "never" to "sometimes but not often" re "encountering dangerous people" as compared to "never" versus "often" (as employed by the investigators).

### Trade-offs, Segments, and Nonlinearity

Whereas ten of the eleven attributes appear to have a "favorable" (e.g., low cost per mile, high level of comfort), and unfavorable end (e.g., high cost per mile, low level of comfort), for most all S's, one

attribute, "opportunity to socialize", may be non-linear in preference levels which contributes to attribute interaction problems perhaps not well handled in the trade-off process, particularly where segments may exist. For example, if one wants to "often have an opportunity to socialize" then "longer total travel time" may be desirable, but only if "the possibility of encountering dangerous people" is "never." Intriguing interaction possibilities are endless. And although the authors recognize this nonlinearity issue they interpret it as a "golden mean" rather than as conflicting segments at the extremes of the dimension. Thus, as they note, not only is multiple regression questionable, but so is any linear assumption model.

It is equally important to the interpretation of these data to note that trade-offs were made for transportation modes in general, not for specific modes. Does this procedure result in different attribute importance inferences than if specific mode comparisons (e.g., auto-bus, auto-train, auto-carpool, etc.) had been employed? Again, attribute interaction combined with "situational" interaction (mode as situation, work versus school as situation, "in hurry" versus "not in hurry" as situation, etc.) seem relevant here.

A Study of Public Policy Impact  
on Consumer Decision-Making  
Philip C. Burger and Alladi Venkatesh

### Limitations

The study as reported leads to the impression that it was initially undertaken to describe and "segment" the CB user market. The transition between these analyses and the simulation referred to is ambiguous. Were specific belief and attitude semantic differential items dealing with "telephonic needs" and "entertainment" included, and how could a simulation of the effect of the proposed change(s) be conducted without any basis for estimating the nature and degree of perceptual (attitude-belief) change? Would not exposing a sample of respondents within each of the key segments to the proposed change at least in concept description form have provided helpful information for the simulation? And, exactly how does one compare the value of the increase in satisfaction with telephonic communication with the value of the decrease in satisfaction in entertainment value... how was satisfaction measured?

### Contribution

A research contribution to public policy decisions is always laudable, even if the mechanics of the application are not always clear. Further, it seems particularly instructive that two rather different segments were identified here with respect to primary satisfactions sought in the use of the product; indeed, not only different, but antithetical in terms of the projected impact of the policy decision. This may reasonably be expected to be the case in various other policy decision contexts, and yet is rarely noted in discussion or research. For example, if younger children use advertising differently than older children, does the correct policy decision "trade-off" the sizes of these segments taking into consideration the costs and/or benefits associated with each proposed policy action? Or, is it even reasonable to talk about "tradeoffs" in comparing segments for policy decision purposes?

### Directions for Future Research

Historically, it may be true that when consumer researchers "discover" a new area of inquiry, in this

case, public policy applications of consumer research, there is some tendency to let tool or technique considerations gain more attention than they should, and in so doing, sometimes to mask more fundamental issues. Each of use has our own favorite technique. But can't we let the problem determine the technique rather than vice-versa? Are utilities for CB and transportation mode utilities best determined via regression weights or trade-off analysis derived part-worths? Or doesn't it make any difference?

And, individual versus aggregate levels of analysis continue to pose interpretation problems. Certainly, it makes sense to retain individual level data if we are going to estimate the number of people who would shift this way or that given some change in marketing mix variables.

With respect to these issues, it would be desirable to have a "technical conference" co-sponsored by ACR and by all the Federal Agencies and other regulatory or quasi-regulatory groups so that standards regarding types of analyses, general design issues, etc., could be developed.

Two other suggestions seem pertinent:

(1) Most any research in public policy applications of consumer behavior seems unavoidably concerned with consumer satisfaction/dissatisfaction issues. Standardization in definition and measurement of CS/D dimensions would be very desirable so that "normative" or benchmark data could be developed in the various domains of public concern with consumer behavior. How "satisfied" (on a commonly employed scale) does a person have to be in order to ride a bus versus a car? Is there a comparable level of satisfaction required before a person will switch to a lower energy usage form of transportation? Should we be more concerned with the satisfaction enhancement or the dissatisfaction minimization side of the policy implication question?

(2) Just as the FTC and presumably other agencies have determined the worthiness of establishing priorities on issues to be investigated, so might ACR'ers be guided by such a schema. As consumer researchers, should the dollar impact (cost/savings) of a policy decision on consumers be considered as importantly as the increase in satisfaction or the years of additional health or X utilities of perceptions of increased personal competence in the marketplace? If we do not become entangled with "the criterion question" when we see regulators perhaps less prepared than we not fearing to plunge in, do we do a disservice to our public constituents?



Frederick D. Sturdivant, The Ohio State University

### Introduction

Marketing seems to have suffered throughout most of its history as an academic discipline with an inferiority complex. At first it was probably related to our standing relative to the parent discipline of economics. Not only do they not give a Nobel prize in marketing, but marketing scholars have yet to come up with a label which enjoys such universal usage and status as "economist." It would appear that little improvement in the discipline's mental health has accompanied its expanding identification with other basic disciplines such as mathematics and psychology.

Perhaps one of the major burdens of the field has been a rather deep-rooted need to have a practical effect. While researchers in other disciplines seem not to be terribly troubled by relevance, usefulness, and the like, marketing scholars appear to have a more empirical-practical bent. And yet, the results on that score are not very encouraging. As I wrote in the Preface to Managerial Analysis in Marketing:

The objective observer would have to admit that most of the significant conceptualizations and applications have originated with practitioners in the field. In like manner, scholars of marketing have played only a minor role in the formulation of public policy (Sturdivant et al., 1970).

Given our misgivings about lack of impact, it is not surprising that the three papers included in this session received positive votes from the reviewers. Two papers not only focused on "real" problems, but were empirically based and, best of all, were submitted to agencies of the federal government (Department of Transportation and the Federal Communications Commission) to guide them in making policy. As a reviewer of one of these papers commented, "[The] paper presents an interesting case or example of consumer research used in public policy." Another reviewer said about one of the papers, "I like it because it was done to assist the development of real-world policy."

### The Two Empirical Studies

The reviewers placed considerable emphasis on the applied nature of the two empirical papers. One reviewer in commenting on the Golden, Betak, and Alpert paper said that it was "An interesting application of conjoint analysis to transportation mode choice. The problem is not new. The methodology isn't either, but applications have not been published very much." While one would assume that I had not been selected as a discussant because of my extensive use of conjoint, discriminant, and regression analyses, the reviewers' emphasis on utility came as a great relief.\*

\*One of my former doctoral students described me in a recent letter as "an individual who wouldn't know a Chi-Square from a Mann-Whitney. Although I will concede that you probably have a thorough familiarity with the Mann Act." I trust his career will flourish at Guatamala Tech.

Given the importance of the usefulness criterion, the Golden et al. paper as well as the one prepared by Burger and Venkatesh will be evaluated largely on that dimension. The latter paper focuses on a problem which grew out of the enormous boom in citizen's band radio usage in this country. While the problem may not rank with the search for a cure for cancer in national priorities, capacity constraints on CB channels and the telephonic/entertainment trade-off issues are by no means trivial in the eyes of some 20 million Americans.

Having conceived an "emergent policy consumer paradigm" and collected and analyzed a variety of data related to CB ownership, usage, and attitudes, one needs to consider the utility of the conclusions reached. It would appear that Burger and Venkatesh offer some mildly supportive words for the notion of increasing the number of radio bands and hence the telephonic properties of CB radios. The lack of strength in their recommendations may be attributed to the broad scope of their research activity. In sum, did the factor analysis of psychographic data or the semantic differential analysis provide data which were directly related to the research question? My principal concern about the paper was the lack of specificity in research design and conclusions vis-a-vis the policy issue. It would be interesting to know the extent to which their results influenced the FCC.

The paper authored by Golden, Betak, and Alpert also addresses a question of considerable importance to millions of citizens. Indeed, the reluctance of Americans to use mass transit facilities has been a subject of considerable interest to urban planners and transportation policy-makers for years. As noted above by one of the paper's reviewers, the authors for the first time applied conjoint analysis to the problem of attracting potential mass transit users.

While one might express concern about the reasonableness of such variables as "cost per mile" with a range of thirty cents, and the authors' assumption of respondent rationality as the basis for their conclusions, perhaps the principal focus again should be the usefulness of the findings. There would seem to be little question that the paper has made a contribution simply by exposing some policy-makers to the use of conjoint analysis. The interrelatedness of attributes is clearly revealed. The authors also offer a realistic perspective in noting that it may well be politically and/or economically impossible to alter the critical variables.

The research findings, however, are more exploratory in nature than actionable. As such they may well provide more of a basis for further research than for policy formulation. Given the apparent strength of feeling in favor of applied research, it would be helpful to know how the sponsoring agency did react to the report. Did the policy-makers understand the analysis? Were they inclined to base decisions on the results? While the authors should not suffer any great pangs of guilt over the exploratory nature of their work, they should feel guilty about heaping more jargon on an already jargon-laden world. In the conclusion of their paper they write:

Finally, in conjunction with the recommended

study of incremental changes, further development should be undertaken of more parsimonious instrumentation for eliciting trade-off data from potential users of transportation services, minimizing the respondents' time investment, and reducing the computational costs of analyzing trade-off data.

I think that means that we need a simplified form of data collection and analysis. Given the bureaucrats' affection for obfuscation and their enthusiasm for murdering the English language, Golden and her colleagues need not contribute more twaddle to their vocabulary. This paper deserves to be quoted in Edwin Newman's next book!

In sum, both of these papers applied relatively sophisticated analytical techniques to two rather important problems. Their efforts are significant and should be applauded. In both instances there appears to be a need for greater precision and clarity in presenting the results to the general readership and to those who are formulating policy in these areas. They should also recognize that if more applied research is going to be undertaken and if that research is to be productive, writers will need to share their application results more fully. They need to critique their own performance vis-a-vis the utility of their research. Without some discussion of client reaction, whether it be positive or negative, objective or political, other researchers will gain little from their real-world experience. Indeed, one of the hazards of failing to treat application experiences is that it may be tempting for some to fit a few exotic research methods to a "hot" policy problem without any concern for results.

#### A Measure for Market Delineation

While the two empirical papers treated topics of considerable importance, the conceptual paper by Bourgeois considers one of the most vital areas of public policy--antitrust. One would think that having been in the antitrust business for almost a century, the courts would have come up with a reasonably consistent measure of market delineation. Such is not the case, however. The fundamental and essential notion of relevant market is the subject of much confusion and considerable inconsistency.

Brown Shoe Company v. United States has long served as a bench mark for the discussion of relevant market. While it is not our purpose to present a review of the cases in this area, a few examples may serve to illustrate the complexities of the topic addressed by Bourgeois in his paper. In Brown Shoe the Supreme Court ruled that:

The outer boundaries of a product market are determined by the reasonable interchangeability of use or the cross-elasticity of demand between the product itself and substitutes for it. However, within this broad market, well-defined sub-markets may exist which, in themselves, constitute product markets for antitrust purposes. The boundaries of such a sub-market may be determined by examining such practical indicia as industry or public recognition of the sub-market as a separate economic entity, the products' peculiar characteristics and uses, unique production facilities, distinct customers, distinct prices, sensitivity to price changes, and specialized vendors (Brown Shoe Co. v. United States, 1962).

One might conclude that a product demonstrated to enjoy a high degree of interchangeability with other products would have a relatively broadly defined relevant market. Nevertheless, in a 1976 decision in the Eastern District of the U.S. District Court in Pennsylvania, the court ruled against the Mrs. Smith's Pie Company even though consumer research established a high degree of substitutability between frozen pies and other desserts. In describing the research conducted by Professor Daniel McLaughlin, the court wrote:

Consumers were given a list of food items and told to divide the list into groups in terms of menu planning and use. . . . More than 95 percent of those interviewed perceived frozen dessert pies as interchangeable with at least some other dessert item, and more than 90 percent actually would substitute another item for frozen pie (United States of America v. Mrs. Smith's Pie Co., 1976).

Even though the court rejected the testimony of the government's expert witness as "completely useless," the court ruled in favor of the government:

[The] government contends that frozen dessert pies constitute a distinct sub-market within the broader market encompassing all desserts, and that the sub-market is itself a "line of commerce" for the purposes of the Clayton Act. We agree.

This decision would appear to run counter to earlier rulings such as United States v. E. I. du Pont de Nemours & Co. which in 1956 emphasized "reasonably interchangeable" as the guideline for relevant market. The recent ruling in favor of IBM over Peripherals Leasing Corporation and Memorex Corporation relied heavily on du Pont-like standard. Hence, the Bourgeois paper addresses a topic greatly in need of research and clarification.\*

Bourgeois' general concept of the market as a complex set of relationships between producers and consumers is sound. His emphasis on consumer perception of products is to be applauded. His assertion, however, that "The output on the supply side is clear--a product(s)" is simply not correct. For example, what is the relevant market for a supermarket with several thousand items ranging from traditional grocery, meat, and produce items to garden supplies and greeting cards? More directly, what is the product the consumer perceives: a cluster of goods and services, a market basket, convenience?

It is in this context that Bourgeois' model for market delineation may be viewed. From a mechanical standpoint his heuristic is useful and workable. It would be enormously time-consuming in a multiple product setting, but it could offer a degree of precision often missing from such estimates. As suggested by the supermarket example, however, one should not assume that "product" is that easily defined. Nor, it should be noted given the Mrs. Smith's Pie Company case, should one necessarily assume that courts will accept consumer perceptions as the basis for ruling on relevant market.

---

\*For a detailed and thoughtful discussion of this topic, see George S. Day, William F. Massy, and Allen D. Shocker, "The Public Policy Context of the Relevant Market Question," Working Paper 221, Graduate School of Business, University of Pittsburgh.

Bourgeois has indicated that application of his framework and methodology is the next task. Such a step might be premature. A method is not going to resolve the definitional problems surrounding the relevant market concept. The inconsistencies in the rulings between merger and non-merger cases, between various district courts and circuits, suggest a degree of conceptual confusion which is unlikely to be corrected simply by a research tool. A thorough review of court decisions, law review articles, literature on market segments, and perhaps a number of other topics will have to be explored in hope of constructing a theory of relevant market. In many respects, the effort may well parallel the contribution Areeda and Turner (1975) made to the concept of predatory pricing. One would hope, however, that it would be less dependent on traditional micro-economic theory than was true in the Areeda and Turner effort. In sum, as a critical element of antitrust, relevant market definition is of vital importance and deserves a thorough and broadly-based research effort.

#### Conclusion

A certain satisfaction may be derived from the fact that marketing scholars are being called on to contribute to the formulation of policy. The relevance of research conducted in the field can only be improved by such working relationships. Application does serve as an important guard against trivial and otherwise meaningless research. Application per se does not, however, necessarily serve as a barrier to poorly conceived research or poor communication of findings. It is also likely that the more broadly-based and theoretically-sound the research, the more likely it is to have a positive reception. "Quick and dirty" applications should not flourish. Perhaps the best measure marketing scholars will have in terms of quality will be the implementation of their recommendations with good results. If such efforts are to have a full impact, the application aspects must be shared with other researchers. Marketing needs to establish a tradition of reporting on the reaction of research users to the efforts of marketing scholars.

#### References

Phillip Areeda and Donald F. Turner, "Predatory Pricing: Related Practices Under Section 2 of the Sherman Act," Harvard Law Review, Vol 88:697 (1975), pp. 697-733.

Brown Shoe Co. v. United States, 370 U.S. 294, 82 S. Ct. 1508, 8 L. Ed. 2d 510 (1962).

Frederick D. Sturdivant et al., Managerial Analysis in Marketing (Glenview, Illinois: Scott, Foresman & Co., 1970).

United States of America v. Mrs. Smith's Pie Company, 440 Federal Supplement 220, 227 (1976).

SURVEY DATA RELIABILITY EFFECTS ON  
RESULTS OF CONSUMER PREFERENCE ANALYSES

Abraham D. Horowitz, General Motors Research Laboratories  
Thomas F. Golob, General Motors Research Laboratories

Abstract

Mail panel survey respondents were resurveyed to assess data reliability. Aggregate measures of evaluations toward hypothetical new products were found to be reliable. However, certain respondents were unreliable, and decision criteria were developed for identifying such respondents when resurvey data are unavailable. The effects on data analyses results of eliminating potentially unreliable respondents from the survey sample were investigated. The magnitude of the effects depended upon the type of analyses conducted.

Introduction

Reliability is an essential characteristic of a measuring instrument. Synonymous for reliability are dependability, stability, consistency, and predictability. A measure is said to be reliable if it produces the same results on different occasions when conditions are kept constant. Reliability is a necessary but not sufficient condition for validity.

The concept of reliability in social sciences was developed in the context of psychological and educational measurement and was focused on how consistently skills of individuals are rated on two successive occasions (test-retest procedure). If the time interval between test and retest is short enough, changes can be attributed to the unreliability of the measuring instrument and/or the individual rather than to a true change in the individual. In the clinical and social psychological literature there is abundant evidence that attitude scales, like individuals' skills scales, are reliable, yielding comparable results when administered on different occasions (Shaw and Wright, 1967).

In consumer behavior literature, on the other hand, the issue of reliability is seldom discussed, and data regarding the reliability of the instruments used in consumer research are rarely provided. As an illustration, Jacoby (1975) reviewed 300 brand loyalty studies finding only three studies which measured data test-retest reliability. As a second illustration, at the previous two annual meetings of the Association for Consumer Research (1976 and 1977) apparently only two papers dealt with data reliability (Peter, 1977 and Best et al., 1977). The first study discussed the methodological problems encountered in the measurement of data reliability and the second study reported high reliability at the aggregate level, but low to moderate reliability at the individual level. The present research is aimed at explaining such distinctions between aggregate and individual level results.

The purpose of the present research is twofold. The first objective is to test the reliability of consumers' evaluations of hypothetical new personal transportation vehicles elicited through a mail panel survey. This is accomplished through application of a test-retest procedure. The second objective is to test whether elimination of potentially unreliable survey respondents significantly affects results concerning consumers'

preferences among the hypothetical concepts and attribute importances inferred from choice models. This is accomplished through comparison of results from analyses of the entire mail-panel survey sample with those from analyses of subsamples. These subsamples are determined by application of criteria developed in the test-retest analyses to identify potentially unreliable respondents.

Data

A nationwide mail panel survey of 1565 consumers was administered to elicit consumer responses to four hypothetical new transportation concepts. For comparison, evaluations of the present vehicle driven or ridden in most often by respondents were also collected. Concept presentations consisted of words and sketches. In recognition of the complexity of the survey, each respondent was presented with only three of four concepts, chosen and ordered randomly.

In order to test the reliability of these data, fifty-six respondents were randomly chosen from the initial survey and were asked to respond again to the same questionnaire four months after the full scale survey. Survey responses of the fifty-six respondents on the initial surveys are referred to as "test" data, while responses to the same questions collected in the follow-up survey are referred to as "retest" data.

Respondents were asked first to rate how much they like or dislike each hypothetical concept (called hereafter, 1st affect scale). Affect ratings were also obtained for the presently owned vehicle which each respondent drove or rode in most often. Then, respondents rated their satisfaction with their present vehicle and with each concept on each of seventeen attributes. Following satisfaction ratings, respondents were asked to rate how much they agree or disagree with nine statements concerning the concepts (called hereafter agree-disagree scales). Finally, respondents were asked again to state their like or dislike with the concept (2nd affect) and then to state their intention to purchase each concept "if this vehicle were available today" (purchase intention scale).

All the scales have seven categories, (-3, -2, -1, 0, 1, 2, 3), with the exception of purchase intention which has five categories. For purposes of interpreting results in the present research, liking, positive intention, satisfaction and agreement are described by positive numbers; dislike, negative intention, dissatisfaction and disagreement are described by negative numbers. Neutral attitudes are represented by the number zero.

## Test-Retest Results

### Aggregate Measures

The differences between test and retest means are small for all scales, their absolute values being typically less than or equal to 0.5, or 1/2 of the interval between two adjacent categories. Average test-retest ratings for each scale are presented in Appendices A and B. None of the differences were significant at the  $\alpha = 0.01$  level for univariate t-tests, suggesting that no systematic shifts in aggregate attitudes occurred. Moreover, correlation coefficients between test and retest computed over all average ratings were very high, ranging from 0.94 to 0.97 for the concepts ( $n = 29$  scales for each concept). The order of the average Affect and Purchase Intention ratings of the concepts and the present vehicle are the same in the test and retest. Therefore, at the aggregate level -- when individual test-retest changes are ignored -- the survey results were very reliable.

### Disaggregate Measures

A different picture emerges when the test ratings for each scale are related to the retest ratings, taking the individual respondent as the unit of observation. Correlations between test and retest for each of the twenty-nine measures (affect, purchase, intention, satisfaction and agree-disagree scales) were calculated for each concept separately. The number of scales with correlations significantly different from zero at the 0.01 level ranged from four (14% of the scales) to twelve (41%) for the concepts and was 13 (72%) for the present vehicle. The significant test-retest correlations were in the moderate range: 0.33 - 0.70. These results imply that a simple linear model is not able to predict the retest data from the initial test data on an individual-by-individual basis for most of the scales.

The insignificant correlations may be due in part to the concentration of responses within a limited range of the scales or to attenuation due to interval assumptions applied to ordinal quality data (Bohrnstead, 1970).

To avoid possible shortcomings of correlation analysis an alternative approach based upon the average absolute

deviation between test and retest was developed.

Defining  $d_{xy} = \sum_{i=1}^n |x_i - y_i| / n$ , where  $n$  is the sample size

and  $x_i$  and  $y_i$  denote respondent  $i$ 's scale responses in the test and retest, respectively, an iterative algorithm was devised to compute probabilities that  $d_{xy}$  will be less than certain values assuming random distributions of responses. The absolute deviation measures computed for each scale and for each concept were then compared to the 0.01 level critical values found for  $d_{xy}$ . The responses on most of the scales (21, 23, 24 and 28 out of 29 scales) for the four concepts and on all scales for the present vehicle were found to be significantly different from random data at the .01 level.

Thus, while test-retest correlations computed on an individual-by-individual basis are not significantly different from zero for most of the scales and are low to moderate for the other scales, hypotheses of random test-retest relationships are consistently rejected.

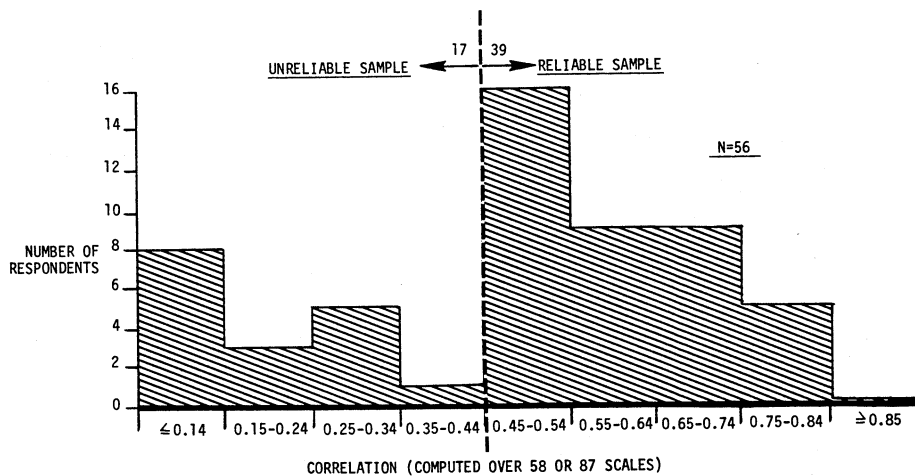
### Characteristics of Unreliable Respondents

It was hypothesized that the low to moderate correlations were due to the inclusion in the sample of certain unreliable respondents. In order to identify such respondents a correlation coefficient was computed for each one of the 56 individuals in the sample. For a respondent with complete data, the sample size for this computation was 87 scales (29 scales for each of three concepts). Figure 1 displays the distribution of individual correlations. Since the distribution is bimodal, two types of respondents emerged: "reliable respondents" characterized by moderate and high correlations ranging from 0.45 to 0.79, and "unreliable respondents" by low correlations from -0.16 to 0.38. Remarkably, all 39 correlations greater than 0.45 were significantly different from zero at the  $\alpha = 0.001$  level, whereas the other 17 correlations were not.

Two criteria were found which effectively distinguished between reliable and unreliable respondents using data collected at only one point in time:

FIGURE 1

TEST-RETEST INDIVIDUAL CORRELATIONS DISTRIBUTION



1. The proportion of the respondent's total scale answers which were in the most frequent scale category for that respondent (i.e., the respondent's propensity to provide the same answer for every question).
2. The proportion of unanswered scales (missing data).

Figure 2 and 3 display the distributions of these two criteria. The first figure shows that among the unreliable respondents there was a considerably greater tendency to provide the same answer for every scale, that is, to mark the same scale category. For example, while 25% of the unreliable respondents marked the same scale category in more than 50% of their responses, only 3% of the reliable respondents did so.

FIGURE 2  
DISTRIBUTION OF PROPORTION OF RESPONSES  
IN MOST FREQUENT CATEGORY

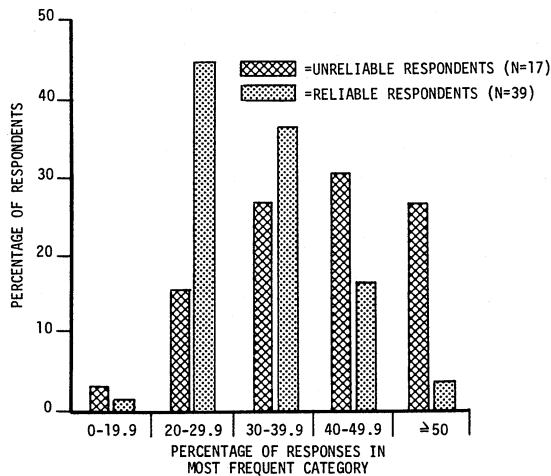


Figure 3 shows that while almost 70% of the reliable respondents answered all the scales, 65% of the unreliable respondents skipped some of the scales.

The two criteria can be used to identify potentially unreliable respondents in a particular data set by establishing a cutoff point for each criterion. Table 1 shows that if these cutoff points are set for the test data such that only those respondents in the test whose most frequent rated category is less than 40% and the percentage of missing data is less than 6% are retained, then the rate of correct classification of reliable respondents is 79% and of unreliable respondents is 76%. The optimum cutoff point must be decided conditionally upon the availability of data and the cost involved in discarding data.

As expected, the elimination of unreliable respondents leads to an increase in the number of scales with correlations significantly different from zero; for the total sample 30% of all scales had significant correlations, which increased to 48% for the reliable subsample. Moreover, test-retest absolute deviations were computed for all scale and concept combinations for the reliable subsample. The number of deviations that are smaller than the critical value for the 0.01 level increased from 85% to 92%.

Application of Reliability Criteria

Subsample Determination

For each respondent to the nationwide mail-panel survey (n = 1565) two measures were computed based upon the test-retest results: (1) proportion of unanswered scales (i.e., missing data) and (2) proportion of responses in the most frequent scale category.

FIGURE 3  
DISTRIBUTION OF PROPORTION OF MISSING DATA

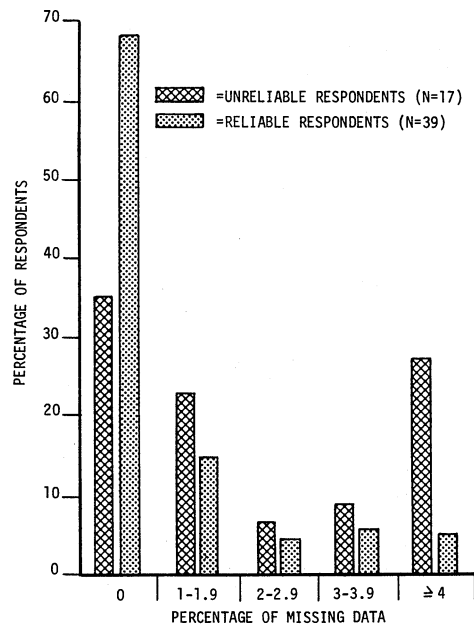


TABLE 1

A DECISION MATRIX FOR ELIMINATION  
OF UNRELIABLE RESPONDENTS

		TRUE STATE (based on test-retest correlation)	
		RELIABLE ( $r \geq 0.45$ )	UNRELIABLE ( $r < 0.45$ )
DECISION* (based on test only)	RELIABLE	HIT 31 (79%)	MISS 4 (24%)
	UNRELIABLE	FALSE ALARM 8 (21%)	CORRECT REJECTION 13 (76%)
	TOTAL	39 (100%)	17 (100%)

\*DECISION CRITERIA:

- A RESPONDENT IS RELIABLE IF AND ONLY IF:
- (1) PROPORTION OF MOST FREQUENT RESPONSE CATEGORY IS LESS THAN 40%, AND
  - (2) PROPORTION OF MISSING DATA IS LESS THAN 6%

Figures 4 and 5 present the frequency distributions of these two measures. Figure 4 shows that a substantial portion of the sample had a rate of missing data of more than 1%. Figure 5 shows that for many respondents (16%) the proportion of responses in their most frequent category was between 40% and 50%. Based then on the shape of these two distributions, and the decision rules developed for the test-retest sample (Table 1), two levels of acceptance for missing data and two levels for the most frequent response category were chosen. These two levels are shown in Table 2. For a given cell in the matrix of Table 2, a respondent who did not pass both acceptance levels was identified as being unreliable and was eliminated from consideration. The sizes of the remaining subsamples ranged from 41.6% to 76.5% of the total sample. This range includes the percentage of test-retest respondents found in the test-retest to be reliable (70%).

#### Attitudes Toward Radically New Concepts

It is assumed that unreliable respondents, when presented with hypothetical new products or services, respond in random fashion or tend to provide the same answer for all questions. Consequently, the average of their responses for a given scale should approximate zero, the scale's middle ground. This leads to the first hypothesis regarding the effect of eliminating potentially unreliable survey respondents.

#### Sociodemographic Characteristics of Potentially Unreliable Respondents

Results summarized in Table 3 show that respondents who satisfy the two reliability criteria are younger and better educated than those who do not satisfy the criteria. These differences are statistically significant at the  $\alpha = 0.01$  level. No statistically significant differences among the subsamples and the total sample were found with respect to income, sex, household size, auto ownership or housing type.

FIGURE 4

DISTRIBUTION OF PROPORTION OF MISSING DATA

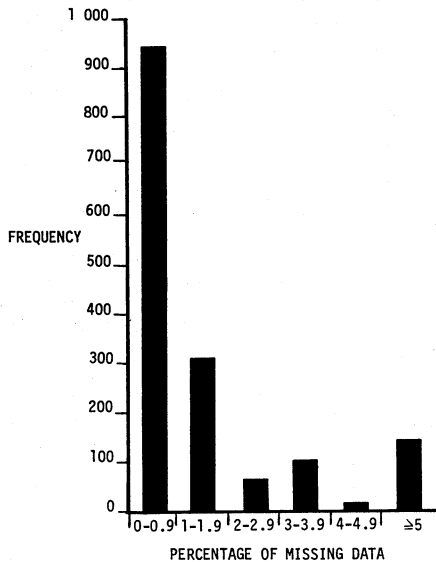


FIGURE 5

DISTRIBUTION OF PROPORTION OF RESPONSES IN MOST FREQUENT CATEGORY

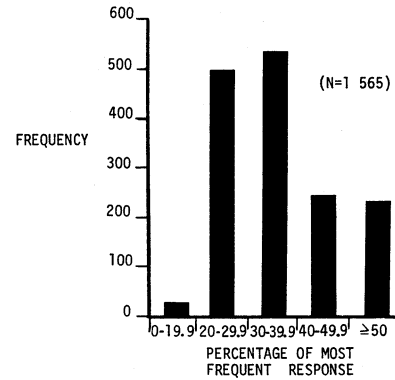


TABLE 2

FOUR SUBSAMPLES DETERMINED BY APPLICATION OF TWO LEVELS OF EACH OF TWO RELIABILITY CRITERIA

		ACCEPTABLE PROPORTION OF MISSING DATA	
		≤1%	≤5%
ACCEPTABLE PROPORTION OF RESPONSES IN MOST FREQUENT SCALE CATEGORY	≤40%	651 (41.6%)	990 (63.3%)
	≤50%	781 (49.9%)	1 198 (76.5%)

(N=1 565)

TABLE 3

STATISTICALLY SIGNIFICANT SOCIODEMOGRAPHIC DIFFERENCES AMONG SUBSAMPLES AND THE TOTAL SAMPLE

SOCIODEMOGRAPHIC CHARACTERISTIC	TOTAL SAMPLE (N=1 565)	SUBSAMPLES DEFINED BY RELIABILITY CRITERIA			
		5/50 (N=1 198)	5/40 (N=990)	1/50 (N=781)	1/40 (N=651)
AGE OF RESPONDENT					
MEAN (YEARS)	40.6	38.8	38.3	37.4	36.9
STANDARD DEVIATION	13.9	13.1	13.0	12.2	12.0
PROPORTION OF RESPONDENTS WHO ATTENDED COLLEGE	0.57	0.59	0.60	0.62	0.64

Hypothesis I. The elimination of potentially unreliable respondents to a consumer survey results in a systematic shift in ratings toward hypothetical new products or services on each attitudinal scale:

- (1) An increase in the absolute value of the average attitudinal measure (i.e., a shift in the average rating outward toward the nearest endpoint of the scale); and
- (2) A decrease in data dispersion.

Three of the four hypothetical transportation concepts were radically different from anything commonly used today. Average ratings and standard deviations were calculated for all 29 scales for each of these three concepts. Differences between each subsample and the total sample on the absolute values of the average ratings and the standard deviations were then computed. These data show that the first part of Hypothesis I is confirmed for 22 to 25 scales for one typical concept, depending on subsample. The systematic shift toward the extreme ends of the scales cannot be explained on the basis of pure chance, because the probability of obtaining 21 or more changes in the predicted direction, of the 29 scales, is less than 0.01. Results for the other two concepts are similar, with confirmations for 27 to 29 scales and 21 to 23 scales respectively. The second part of the hypothesis regarding a decrease in data dispersion was confirmed for all scales in all concepts without exception.

Differences among the four subsamples were ordered in the expected direction for each concept. That is, the more severe the criteria of reliability, the larger the differences in both mean and standard deviation between the subsample and the total sample.

#### Attitudes Toward Existing Concepts

Differences in scale means and standard deviations between the total sample and each of the subsamples in respondent's evaluations of their existing vehicles were also calculated. Here Hypothesis I was consistently rejected. All mean ratings for the subsamples are lower in absolute value than those for the total sample, and most standard deviations are higher. Potentially unreliable respondents were more extreme and positive toward their existing vehicles than were the remaining respondents.

Since potentially unreliable respondents were older and less educated than the average respondent, it follows that these respondents are also more satisfied with their present vehicles. This result is consistent with findings (Campbell et al., 1976) that the higher the educational level of an individual the less happy he or she will be with the quality of life, due to higher expectations. A second hypothesis thus emerges.

Hypothesis II. The elimination of potentially unreliable respondents to a consumer survey results in a systematic shift in ratings of well known and strongly liked products or services on each attitudinal scale:

A decrease in the absolute value of the average attitudinal measure (i.e., a shift in the average rating away from the positive end point of the scale).

#### Attitudes Toward Evolutionary New Concepts

One hypothetical transportation concept was considerably more similar to existing automobiles than was any other concept. With regard to this concept, both Hypothesis I and Hypothesis II hold to some degree. Since these two

hypotheses lead to opposite shifts in attitudinal measures, a substantially lesser degree of systematic shift is detectable for this concept. Consequently, a third hypothesis is proposed to cover this composite situation in which neither a totally new product nor an existing one is the subject of consumer evaluations.

Hypothesis III. The elimination of potentially unreliable respondents results in no systematic shift in attitudinal ratings of new products or services which are evolutions or variations of well known existing products or services, due to the joint effects of Hypothesis I and II.

#### Aggregate Preferences for the Concepts

Consumers' aggregate preferences among the concepts were computed from responses to a survey question eliciting a ranking of the concepts in terms of ownership preferences. These rankings are closely related to the order of the concepts in terms of "second affect" rating; in 86% of all cases, if one concept was ranked higher than another, it was also rated more liked or disliked (or the same) on the second affect questions. Consequently, Hypothesis I, II, and III, which cover shifts in mean affect ratings, can be cross-validated using aggregate results from the ranking question.

#### Preference Choice Models

Consumers' preferences among the concepts were modeled in terms of consumer satisfactions with various attributes of the concepts. In this way relative importances of the features were estimated. The specific choice model used was the multinomial logit model (Punj and Staelin, 1976). This model is subsumed in the class of models referred to as strict utility models by economists and Bradley-Terry-Luce models by psychologists. (For an overview of these choice models see Luce, 1977.) It is purported herein that the following two characteristics of the multinomial logit model, together with the small magnitude of the shifts described in the previous Sections, will cause the preference model results to remain approximately unaltered when potentially unreliable respondents are eliminated. It was found empirically that this was indeed the case.

First, the logit model, and most other models in the ascribed class of strict utility models, is specified in terms of utility differences between pairs of choice alternatives. Thus, the specification is unique only up to differences in satisfaction ratings between two concepts on any attribute. Since the shifts in attribute ratings resulting from elimination of potentially unreliable respondents are in the same direction for all concepts, the small magnitudes of these shifts are further cancelled out in computing differences. This is demonstrated in Table 4, which shows that for only 7 of 48 differences across all pairs of concepts for two attributes are there significant differences between a subsample and the total sample at the  $\alpha = 0.01$  level. These two attributes were found to best represent two orthogonal factors determined in principal component analyses of the seventeen satisfaction scales. Results similar to those in Table 4 were found for all remaining attributes as well.

Second, the technique used to estimate coefficients of the utility function specified in the multinomial logit model is that of maximum likelihood. Maximum likelihood estimators for the logit model were shown by McFadden (1973) to be asymptotically unbiased, efficient and consistent. Such desirable statistical properties insure that such estimators are well behaved in the presence of additive disturbances which are independently distributed across the population. A necessary condition for estimator invariance in light



of the introduction of disturbances attributable to unreliable respondents is that such disturbances exhibit the same central tendency as disturbances attributable to other sources, such as noise in measurement and misspecification of utility.

Table 5 shows that the models yield approximately the same results. Thus, inclusion of potentially unreliable survey respondents in this application apparently does not bias conclusions regarding the relative importances of various design features of the transportation concepts. Similar stability would be expected for any of the so-called multi-attribute attitude models frequently used in consumer research. These models (reviewed by Wilkie and Pessemier, 1973) typically hypothesize that a consumer compares products by combining values or utilities on individual features in a manner similar to that underlying the logit model.

Thus, while the shifts in attitudes resulting from removal of potentially unreliable respondents are systematic, since these shifts are small, and since only the relative values of attitudes on different features and different products are used, there are no major changes in preference model results.

TABLE 5

LOGIT CHOICE MODEL RESULTS FOR TOTAL SAMPLE AND SUBSAMPLES

MODEL PARAMETER	TOTAL SAMPLE (N=1 565)	SUBSAMPLES DEFINED BY RELIABILITY CRITERIA			
		5/50 (N=1 198)	5/40 (N=990)	1/50 (N=781)	1/40 (N=651)
GOODNESS-OF-FIT INDICES:					
PSEUDO R <sup>2</sup>	0.28	0.24	0.29	0.28	0.29
% CORRECT CLASSIFICATION	81.3	77.9	80.6	80.7	81.4
t-STATISTICS:					
FUEL ECONOMY	-4.22	-3.94	-4.43	-4.57	-4.68
VEHICLE SIZE	-7.78	-6.92	-7.99	-7.99	-7.81
CONSTANT	9.66	9.36	10.07	9.99	10.36
ELASTICITY RATIOS:					
(VEHICLE SIZE/FUEL ECONOMY)					
CONCEPT D	3.17	3.06	3.17	3.24	2.90
CONCEPT C	4.10	3.80	4.08	4.09	3.68
CONCEPT A	3.00	3.00	3.09	3.00	2.83
CONCEPT B	3.08	2.78	2.81	2.81	2.70

TABLE 4

DIFFERENCES BETWEEN CONCEPT RATINGS ON VEHICLE SIZE AND FUEL ECONOMY FOR ALL PAIRS OF CONCEPTS FOR TOTAL SAMPLE AND SUBSAMPLES

ATTRIBUTE SATISFACTION DIFFERENCES BY CONCEPT PAIR	TOTAL SAMPLE (N=1 565)	SUBSAMPLES DEFINED BY RELIABILITY CRITERIA			
		5/50 (N=1 198)	5/40 (N=990)	1/50 (N=781)	1/40 (N=651)
CONCEPT A-CONCEPT B FUEL ECONOMY	-0.36	-0.31	-0.30	-0.32	-0.30
VEHICLE SIZE	-0.51	-0.52	-0.48	-0.52	-0.52
CONCEPT A-CONCEPT C FUEL ECONOMY	0.03	-0.03	-0.07	-0.04	-0.08
VEHICLE SIZE	0.99	1.06	1.02	1.03	1.02
CONCEPT A-CONCEPT D FUEL ECONOMY	0.60	0.56	0.50	0.54	0.48
VEHICLE SIZE	0.39	0.53*	0.50*	0.55*	0.54*
CONCEPT B-CONCEPT C FUEL ECONOMY	-0.33	-0.34	-0.37	-0.36	-0.38
VEHICLE SIZE	0.48	0.50	0.54	0.56	0.61
CONCEPT B-CONCEPT D FUEL ECONOMY	0.24	0.25	0.20	0.22	0.18
VEHICLE SIZE	0.88	1.01*	0.98	1.03*	1.02
CONCEPT C-CONCEPT D FUEL ECONOMY	0.57	0.59	0.57	0.58	0.56
VEHICLE SIZE	0.40	0.51*	0.44	0.47	0.41

\*DENOTES DIFFERENCE SIGNIFICANTLY DIFFERENT FROM TOTAL SAMPLE DIFFERENCE AT  $\alpha=0.01$  LEVEL

References

R. J. Best, D. I. Hawkins, and G. Albaum, "Reliability of Measured Beliefs in Consumer Research," in W. D. Perreault (Ed.), Advances in Consumer Research, 4(1977), 19-23.

G. W. Bohrnstead, "Reliability and Validity Assessment in Attitude Measurement," in G. F. Summers (Ed.), Attitude Measurement (Chicago: Rand McNally, 1970), 80-99.

A. Campbell, P. E. Converse, and W. L. Rodgers, The Quality of American Life (New York: Russell Sage Foundation, 1976).

J. Jacoby, "Consumer Research: Telling It Like It Is," in B. B. Anderson (Ed.), Advances in Consumer Research, 3(1976), 1-11.

R. D. Luce, "The Choice Axiom after Twenty Years," Journal of Mathematical Psychology, 15(1977), 215-233.

D. McFadden, "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka (Ed.), Frontiers in Econometrics (New York: Academic Press, 1973).

J. P. Peter, "Reliability, Generalizability and Consumer Behavior," in W. D. Perreault (Ed.), Advances in Consumer Research, 4(1977), 394-400.

G. N. Punj and R. Staelin, "A Model of the College Choice Process," in K. L. Bernhardt (Ed.), Marketing 1776-1976 and Beyond, 1976 Educators' Proceedings, American Marketing Association (1976), 227-241.

M. E. Shaw and J. M. Wright, Scales for the Measurement of Attitudes (New York: McGraw-Hill, 1967).

W. L. Wilkie and E. A. Pessemier, "Issues in Marketing's Use of Multi-Attribute Attitude Models," Journal of Marketing Research, 10(1973), 428-441.

APPENDIX A

TEST-RETEST AVERAGE AFFECT, INTENTION AND SATISFACTION RATINGS

SCALE	PRESENT VEHICLE			CONCEPT A			CONCEPT B			CONCEPT C			CONCEPT D		
	(1) TEST	(2) RETEST	(2)-(1)	(1) TEST	(2) RETEST	(2)-(1)	(1) TEST	(2) RETEST	(2)-(1)	(1) TEST	(2) RETEST	(2)-(1)	(1) TEST	(2) RETEST	(2)-(1)
1ST AFFECT	+2.4	+2.5	+0.1	+0.1	+0.7	+0.6	+0.1	-0.1	-0.2	+0.1	-0.1	-0.2	-1.3	-0.7	+0.6
2ND AFFECT	-	-	-	0.0	+0.7	+0.7	0.0	-0.1	-0.1	-0.3	-0.3	0.0	-1.1	-1.0	+0.1
PURCHASE INTENTION	-	-	-	-0.6	-0.5	+0.1	-0.7	-0.9	-0.2	-0.8	-0.9	-0.1	-1.2	-1.2	0.0
SAFETY	+2.1	+2.4	+0.3	+0.5	+0.8	+0.3	0.0	+0.2	+0.2	-0.7	-0.6	+0.1	-0.7	-0.6	+0.1
STABILITY-LOW SPEED	+2.4	+2.5	+0.1	+1.3	+1.5	+0.2	+1.0	+1.1	+0.1	+0.9	+1.0	+0.1	+0.9	+0.5	-0.4
STABILITY-HIGH SPEED	+2.1	+2.3	+0.2	+0.7	+0.9	+0.2	+0.1	+0.3	+0.2	+0.1	+0.2	+0.1	-0.1	-0.2	-0.1
EASE OF PARKING	+1.8	+2.0	+0.2	+2.2	+2.0	-0.2	+2.1	+2.2	+0.1	+2.0	+1.9	-0.1	+2.0	+1.9	-0.1
FUEL ECONOMY	+1.3	+0.8	-0.5	+2.5	+2.2	-0.3	+2.3	+2.4	+0.1	+2.6	+2.4	-0.2	+2.2	+1.9	-0.3
ABILITY TO BE SEEN	+2.0	+1.9	-0.1	+0.4	+0.7	+0.3	+0.1	+0.2	+0.1	-0.5	-0.2	+0.3	-1.0	-0.7	+0.3
CARGO SPACE	+1.9	+1.9	0.0	+0.2	+0.3	+0.1	-0.4	-0.1	+0.3	-0.9	-0.7	+0.2	-1.0	-1.2	-0.2
SEATING COMFORT	+2.2	+1.9	-0.3	+0.2	+0.3	+0.1	-0.2	-0.1	+0.1	-0.3	-0.3	0.0	-0.6	-0.8	-0.2
NUMBER OF PASSENGERS	+2.0	+2.2	+0.2	+0.4	+0.5	+0.1	-0.2	-0.2	0.0	-0.4	-0.6	-0.2	-0.8	-0.9	-0.1
VEHICLE SIZE	+2.1	+2.2	+0.1	+0.4	+0.4	0.0	+0.1	-0.3	-0.4	0.0	-0.4	-0.4	-0.7	-0.8	-0.1
HEATER & RADIO	+2.5	+2.3	-0.2	+1.2	+1.5	+0.3	+0.4	+0.4	0.0	+0.9	+1.1	+0.2	+0.1	+0.6	+0.5
POWER OPTIONS	+2.2	+2.0	-0.2	+1.1	+1.3	+0.2	-0.2	-0.5	-0.3	+0.7	+1.2	+0.5	-0.2	+0.2	+0.4
ACCELERATION	+2.0	+1.8	-0.2	+0.1	+0.6	+0.5	-0.1	-0.1	0.0	+0.8	+0.9	+0.1	-1.0	-0.7	+0.3
PURCHASE PRICE	+1.9	+2.0	+0.1	+0.3	+0.3	0.0	+1.1	+0.8	-0.3	+1.4	+0.8	-0.6	+0.3	+0.3	0.0
RANGE	+2.0	+2.1	+0.1	+1.2	+1.4	+0.2	0.0	+0.1	+0.1	+1.0	+1.0	0.0	-0.8	-0.9	-0.1
FLEXIBILITY OF USE	+2.3	+2.3	0.0	+0.9	+1.2	+0.3	-0.1	+0.2	+0.3	+0.0	+0.7	+0.7	-1.2	-0.7	+0.5
TOP SPEED	+1.7	+1.8	+0.1	+0.2	+0.7	+0.5	-0.2	-0.3	-0.1	+0.6	+1.2	+0.6	-1.2	-1.1	+0.1
MEAN	+2.0	+2.0	0.0	+0.7	+0.9	+0.2	+0.3	+0.3	0.0	+0.4	+0.4	0.0	-0.4	-0.3	+0.1
SAMPLE SIZE	56			33			31			34			36		

APPENDIX B

TEST-RETEST AVERAGE AGREE DISAGREE RATINGS

SCALE	CONCEPT A			CONCEPT B			CONCEPT C			CONCEPT D		
	TEST(1)	RETEST(2)	(2)-(1)	TEST(1)	RETEST(2)	(2)-(1)	TEST(1)	RETEST(2)	(2)-(1)	TEST(1)	RETEST(2)	(2)-(1)
REDUCES TRAFFIC CONGESTION	+0.4	+0.9	+0.5	+0.5	+0.7	+0.2	+0.2	+0.6	+0.4	-0.6	-0.1	+0.5
REDUCES AIR POLLUTION	+0.2	+0.7	+0.5	+2.0	+1.8	-0.2	+0.2	+0.9	+0.7	+1.4	+1.9	+0.5
ⓈLIMITS ARE CONFINING	+0.1	-0.3	-0.4	+1.7	+1.6	-0.1	+0.3	+0.6	+0.3	+2.2	+2.1	-0.1
ⓈUSE TOO RESTRICTIVE	0.0	-0.1	-0.1	+1.6	+1.5	-0.1	+0.6	+1.1	+0.5	+2.2	+2.0	-0.2
WOULD BE PROUD TO OWN	-0.3	+0.2	+0.5	+0.1	0.0	-0.1	0.0	-0.4	-0.4	-0.9	-1.2	-0.3
ⓈTOO EXPENSIVE	+0.5	+0.3	-0.2	+1.1	+0.8	-0.3	-0.2	+0.8	+1.0	+0.4	+0.8	+0.4
SIMILAR TO SECOND CAR	+0.6	+1.2	+0.6	-0.7	-0.8	-0.1	-0.4	-0.5	-0.1	-1.6	-1.4	+0.2
WOULD ENJOY DRIVING	+0.2	+0.5	+0.3	+0.5	+0.3	-0.2	-0.2	-0.2	0.0	-0.8	-0.8	0.0
ⓈNEEDS SPECIAL LANES	-0.4	-0.1	+0.3	+1.3	+1.6	+0.3	+1.2	+1.7	+0.5	+2.2	+2.4	+0.2
MEAN	+0.1	+0.4	+0.2	+0.9	+0.8	-0.1	+0.2	+0.5	+0.3	+0.5	+0.6	+0.1
NUMBER OF RESPONDENTS	33			31			34			36		

ⓈNEGATIVE STATEMENTS

THE EFFECT OF VARYING RESPONSE INTERVALS ON THE  
STABILITY OF FACTOR SOLUTIONS OF RATING SCALE DATA

Roger Best  
University of Arizona

Del I. Hawkins  
Gerald Albaum  
University of Oregon

Abstract

The number of response intervals used with rating scales is generally considered to have a limited impact on the obtained results. A univariate comparison of five interval and continuous interval response formats confirmed this. However, a multivariate (factor analysis) comparison produced significantly different solutions depending on the number of scale intervals used. The implications for the validity of multivariate analysis of rating scale data are discussed.

Introduction

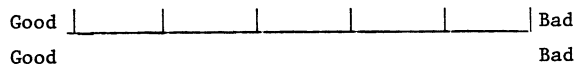
Despite the widespread utilization of rating scales, several issues concerning the appropriate way to utilize these scales remain unresolved. The specific issue to be addressed in this paper is the impact that the number of response intervals has on the results. A number of researchers have addressed this question (Guilford, 1954; Green and Rao, 1970; and Jacoby and Matell, 1971; Lehmann and Hulbert, 1972; Matell and Jacoby, 1972; Masters, 1972; Albaum and Munsinger, 1973; and Bendig, 1974). These studies have focused on univariate comparisons and have generally found that the number of scale positions used has a minimal impact on the obtained results. The authors have recommended using between 3 and 25 intervals depending on the task at hand and the nature of the respondents.

What happens when rating scales using varying numbers of response categories are subjected to a multivariate analysis such as factor analysis? A specific situation can help clarify the importance of this question. At the 1977 Association for Consumer Research conference Vaughn et al (1978) presented an interesting analysis of university choice criteria which utilized several seven-point rating scales as part of its methodology. The responses to these rating scales were analyzed via averaging and by factor analysis. The question now is, "Is it possible that Vaughn et al would have reached different conclusions had they utilized a different number of response intervals.?"

This question is not intended as a criticism of Vaughn, et al. The same question could be addressed to any of the 15 papers at that conference which utilized rating scales and specifically to the 3 which then factor analyzed the results. In utilizing seven intervals, Vaughn et al were following a commonly accepted practice (Tull and Hawkins, 1976). Nonetheless, it is important to determine the possible effects of arbitrarily selecting seven or five or some other number of response intervals.

Method

A five-interval discrete scale which restricts response to one of five intervals and an interally continuous scale (graphic rating scale) which provides unrestricted response were created for use with the semantic differential as shown below.



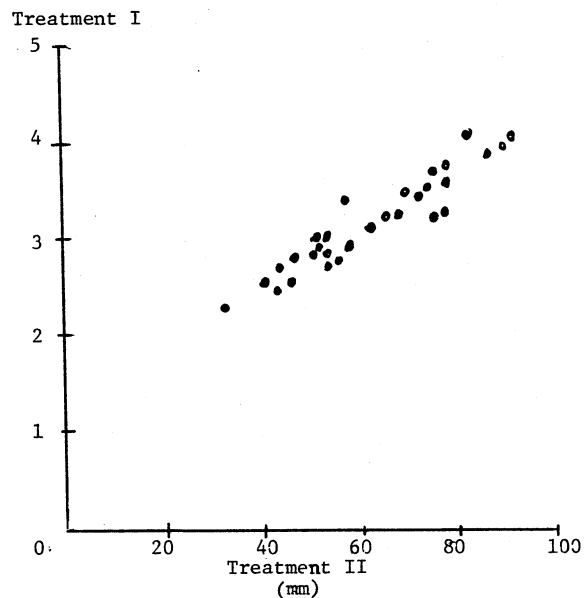
Each scale was 125 mm in total length. A 30-item questionnaire was constructed for each treatment using ten bipolar adjective scales which were used by 176 undergraduate students to evaluate three familiar stimuli--the university, the university bookstore, and the student union. One questionnaire utilized the five-interval discrete scale while the other incorporated the intervally continuous scale. The adjective pairs were selected from among those commonly reported in studies of retail store image that would also be relevant to the stimuli being evaluated in this study. One-half of the students were randomly selected to complete one questionnaire and one-half to complete the other questionnaire. The presentation of scale items and stimuli were randomized in both questionnaires to minimize the chance of order bias.

Analysis and Results

Univariate analysis of the data followed two procedures. First, means were computed for each of the 30 scales (10 scales x 3 concepts) for both the discrete and continuous formats. (The responses to the continuous interval scale were coded using a metric ruler and the values of 1 to 125 mm.). The scale means for the two formats were then compared by correlation analysis. The correlation between the two formats was .92. As Figure 1 illustrates, this represents a very high level of agreement between the two approaches.

FIGURE 1

CORRESPONDENCE BETWEEN SCALE MEANS OBTAINED FROM TWO  
INTERVALLY DISSIMILAR SEMANTIC DIFFERENTIAL SCALES



A second form of univariate analysis was conducted by using the Kolmogorov-Smirnov two sample test. This test is nonparametric and thus removes any concern for the interval nature of this type of data (Labovitz, 1967; Labovitz, 1970; Dawes, 1971; Martilla and Garvey, 1975; Henkel, 1975; and Albaum, et al., 1977). In addition, this test is sensitive to differences in central tendency, dispersion, and skewness. To conduct the test, the continuous scale was converted into five equal intervals that corresponded to the five intervals on the discrete scale. The responses were then assigned the appropriate one to five value.

The results of this analysis are shown in Table 1. Only two significant differences with  $p$  set as high as .15 were found in the 30 comparisons. One would expect this result by chance.

Table 1  
Results of Kolmogorov-Smirnov Test<sup>a</sup>

Semantic Scale	Stimulus Object		
	University	Bookstore	Student Union
Honest	.53	.90	.83
Low quality	.53	1.13 <sup>d</sup>	.60
Low priced	.53	.30	.60
Slow service	.90	.83	.68
Dependable	.75	.75	.60
Clean	.90	.98	.53
Helpful employees	.53	.53	.23
Old-fashioned	.45	.53	1.35 <sup>c</sup>
Unpleasant	.83	.60	.30
Friendly	.53	.23	.15

<sup>a</sup>Z-scores computed from the Kolmogorov-Smirnov test.

<sup>b</sup>Only the left-hand term is shown.

<sup>c</sup>Significant at  $p = .05$ .

<sup>d</sup>Significant at  $p = .15$ .

Based on these two univariate comparisons, the data obtained from these two distinct sets of response categories appear to be equivalent. However, before concluding that there is complete equivalence, we need to take the analysis one step further; we need to examine the equivalence of multivariate analysis of these two scale formats. This is something previous studies have failed to do.

Factor analysis is perhaps the most common multivariate technique applied to rating scale data. Twenty percent of the studies using rating scales in last year's ACR conference performed factor analysis on those scales. To test the sensitivity of factor solutions to varying numbers of response intervals, the responses from each of the two scale treatments for each concept were subjected to normalized and varimax rotated factor analysis. The factor solutions for each concept were compared by computing the correlation between the factor loadings obtained from each treatment type. The correlations for the three concepts ranged from .45 to .55. The nature of these responses are shown for one concept, the University, in Table 2 and illustrated graphically in Figure 2 (of the three concepts this one had the highest degree of linear association between the two factor solutions).

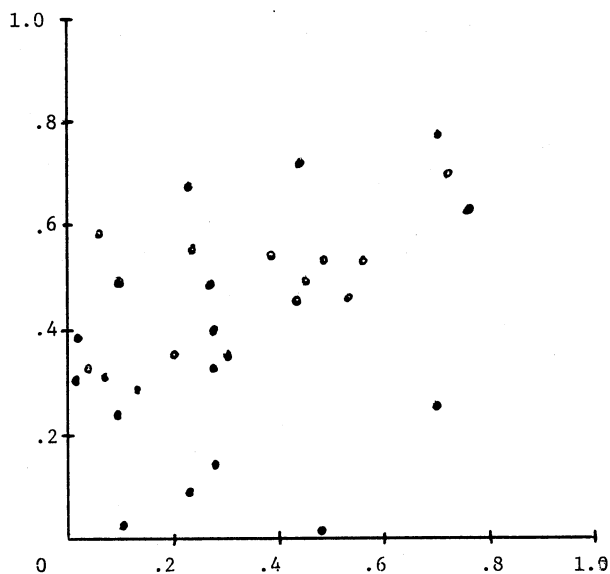
Table 2  
Factor Analytic Solution of  
5-Interval and Continuous  
Semantic Scale Types for "University"<sup>a</sup>

Scale	Factor I		Factor II		Factor III	
	5 inter-val	con-tinu-ous	5 inter-val	con-tinu-ous	5 inter-val	con-tinu-ous
Honest	.39	.70	.34	.32	.47	-.01
Low quality	.26	.66	.38	.04	.59	.10
Low priced	.75	.26	.34	-.21	.22	.09
Slow service	.23	.48	.48	.14	.56	.44
Dependable	.40	.50	.22	.30	.52	.50
Clean	.18	.30	.33	.20	.78	.62
Helpful employees	.57	.25	.43	.45	.42	.53
Old-fashioned	.24	.13	.74	.72	.28	.12
Unpleasant	.33	.10	.70	.68	.30	.40
Friendly	.39	-.03	.10	.02	.43	.43

<sup>a</sup>Correlation between factor loadings is .55.

FIGURE 2

CORRESPONDENCE BETWEEN FACTOR LOADINGS OBTAINED FROM FACTOR ANALYSIS OF FIVE INTERVAL DISCRETE AND CONTINUOUS INTERVAL SEMANTIC DIFFERENTIAL SCALE DATA



Concept ; University;  $r = .55$

## Discussion

The univariate analysis of the responses to a five interval and a continuous interval rating scale indicated that the two approaches produced equivalent results. However, correlations of the factor loadings obtained from each type scale ranged from .45 to .55. Thus, the explained variance between the multivariate solutions ranged from 22 to 30 percent. This low level of association questions the reliability and validity of interpretations derived from factor analyzed rating scales.

Since factor loadings are dependent upon the covariance between items as well as the variance within an item and the variances were not significantly different as shown by the Kolomogorov-Smirnov tests, these results suggest that changes in the number of response intervals influences the covariance between items substantially more than it does the central tendency or variance within an item. Another explanation is that factor analysis is inherently more sensitive to minor shifts in responses than are means or aggregate tests such as the Kolomogrov-Smirnov test.

## Conclusions

In this study, the conclusions one would reach based on the factor analysis would depend upon the number of response intervals used. However, two extreme forms of the rating scale were compared, a five interval discrete scale and a continuous scale. The results indicate that the number of response intervals can influence the factor solutions even though there is no significant variation in their univariate properties. Looking back at Vaughn et al.'s study (as well as all the others that have factor analyzed rating scale data), we can say that it is possible that the factor solutions were influenced by the number of response categories. What must be determined now is how sensitive the procedure is to variations in the number of categories used. Will a shift from 5 to 6, or 5 to 7 influence the result? In addition, one must wonder how sensitive other multivariate techniques such as the principle components derived in discriminant analysis are to changes in the number of response categories. Also, which is more accurate in terms of validity? Perhaps this question can be answered by examining the factor solutions of known interval or ratio scaled data. This data could then be re-measured using a varying number of scale intervals. A comparison of the various solutions would yield some insight into the accuracy of the solutions. Until questions such as these are answered, the results of multivariate analysis, particularly factor analysis, of rating scale data must be viewed with caution.

## References

- Albaum, G., and G. Munsinger, "Methodological Questions Concerning the Use of the Semantic Differential," paper presented at the Southwestern Social Science Association Conference, Dallas, March 1973.
- Albaum, G., R. Best, and D.I. Hawkins, "Measurement Properties of Semantic Scale Data," Journal of the Market Research Society, 19 (1977), 21-28.
- Bendig, A.W., "Reliability and Number of Rating Scale Categories," Journal of Applied Psychology, 38 (1974), 38-40.
- Dawes, R.M., "Suppose We Measured Height with Rating Scales Instead of Rules," Oregon Research Institute Technical Report (1971), 2.
- Green, P.E., and V.R. Rao, "Rating Scales and Information Recovery--How Many Scales and Response Categories to Use?" Journal of Marketing, 34 (1970), 33-39.
- Guilford, J.P., Psychometric Methods, New York: McGraw-Hill, 1954.
- Henkel, R.L., "Part-Whole Correlations and the Treatment of Ordinal and Quasi-Interval Data as Interval Data," Pacific Sociological Review, 18 (1975), 3-26.
- Jacoby, J., and M.S. Matell, "Three-Point Likert Scales are Good Enough," Journal of Marketing Research, 8 (1971), 495-500.
- Labovitz, S., "Some Observations on Measurement and Statistics," Social Forces, 46 (1967), 151-60.
- Labovitz, S., "The Assignment of Numbers to Rank Order Categories," American Sociological Review, 35 (1970), 515-524.
- Lehmann, D.R. and J. Hulbert, "Are Three-Point Scales Always Good Enough?" Journal of Marketing Research, 9 (1972), 444-446.
- Martilla, J.A. and D.W. Carvey, "Four Subtle Sins in Marketing Research," Journal of Marketing, 39 (1975), 8-15.
- Masters, J.R., Reliability as a Function of the Number of Categories of a Summated Rating Scale, unpublished dissertation, University of Pittsburgh, 1972.
- Matell, M.S. and J. Jacoby, "Is There an Optimal Number of Alternatives for Likert-Scale Items," Journal of Applied Psychology, 56 (1972), 506-509.
- Tull, D.S. and D.I. Hawkins, Marketing Research: Meaning, Measurement and Method. New York: Macmillan, 1976, 336.
- Vaughn, R., J. Pitlik, and B. Hansotia, "Understanding University Choice: A Multi-Attribute Approach," in H.K. Hunt, Advances in Consumer Behavior, Vol.V, (Ann Arbor: Associations for Consumer Research, 1978), 26-31.

# THE EFFECTS OF SPONSORSHIP ON MAIL SURVEY RESPONSE AND EVALUATION BIAS

Kenneth E. Miller, University of Utah  
Marilynn Turner (student), University of British Columbia

## Abstract

A field experiment was conducted to test two mail survey sponsorship hypotheses (1) that commercial sponsorship would produce a higher response rate than public sector or independent research agency sponsorship and (2) that commercial sponsorship would bias the results of corporate evaluation responses. Customers of a commercial bank were randomly assigned to treatment groups where identical questionnaires were mailed from three different sponsors. Analysis of variance and discriminant analysis indicated that, in this type of situation commercial sponsorship has a favorable effect on the quantity of the response but no practical effect on the quality of the response.

## Introduction

Although the mail survey method is obviously useful in a wide variety of circumstances, it is particularly well suited to the investigation of consumer satisfaction with products and/or services. Because the names and addresses of the population to be studied are readily available, contact by mail is considerably more efficient than either personal or telephone interviews. Moreover, the mail survey method allows the respondents to reflect leisurely on the service they have received and to offer careful evaluations and criticisms. The depth of the responses on open-ended questions in a mail survey provides the users of the research with much more actionable information than the top-of-mind responses generally elicited by telephone or personal interviews.

Despite its abundant advantages, the mail survey method, when used for customer evaluation/satisfaction surveys, presents the researcher with a serious dilemma. It is commonly believed that revealing the commercial sponsorship of such a study will bias the response to evaluative questions. According to this assumption, customers, sensing a lack of anonymity and wishing to please the sponsor, might withhold critical opinions of the service they have received, thus biasing the results in favor of the sponsor. On the other hand, corporate research sponsored by an outside agency may be less biased but more expensive than an in-house study. However, the consumer who is undoubtedly interested in improving services is more likely to return a questionnaire to the supplier of these services than to an independent research agency. Acceptance of these two hypotheses about the effects of sponsorship on the rate of return and the quality of the responses forces the researcher to make a difficult decision about the sponsorship of commercial research.

## Previous Studies

In their review of the literature, Kanuk and Berenson (1975) reported that very little work had been done on the effects of sponsorship on mail survey response rates or bias. They cited Scott's work (1961) which found that British government sponsorship elicited a higher response rate in Great Britain than non-government sponsorship, and Brunner and Carrol's work (1969) which found that university sponsorship elicited a higher response rate than commercial sponsorship. The authors

discounted a third study, this by Baur (1947), because it failed to support its contention that VA sponsorship of a mail survey to ex-servicemen biased the results.

The existing literature offers little substantiation on the effect of sponsorship on response rate and bias in customer evaluations. Baur lends some support to the assumption that commercial sponsorship would bias the results. Brunner and Carrol's research suggests that, in a blind study, university, rather than commercial, sponsorship would be a better alternative.

As a preliminary step to a comprehensive consumer satisfaction study for a medium-sized bank, a study was designed to test several sponsorship hypotheses:

1. that commercial sponsorship where the respondent is a regular customer would produce a higher response rate than either independent agency sponsorship or university sponsorship;
2. that commercial sponsorship would bias the results of a consumer evaluation study relative to results obtained using agency or university sponsorship.

## Method

Three random samples (Groups A, B and C) were drawn from the personal checking account file of a medium-sized bank operating in the state of Utah. The envelopes and stationery of the respective sponsors were simultaneously mailed to the appropriate treatment group. Group A (n = 172) received a questionnaire from the commercial sponsor (bank) with a cover letter signed by a bank officer. Subjects in Group B (n = 168) received a questionnaire from the local university with a cover letter signed by a Business School professor. The questionnaire sent to Group C (n = 168) was accompanied by a cover letter signed by a consumer researcher at a local independent research agency. This medium-sized research agency was known for presenting public interest survey results on a local TV station. All references to the sponsoring bank were removed from material sent to groups B and C.

The three-page questionnaire asked respondents to rate thirteen aspects of bank service, such as "friendliness of tellers" and "accuracy" etc. on a 6-point scale from "Excellent" to "Very Poor." These thirteen items were synthesized from a previous in-house study where 120 evaluation items were factor analyzed. Respondents were also asked to relate any specific problems or bad experiences they might have had during the past year. Demographic information was also collected.

## Results

### Response Rates

Response rates to the alternative forms of sponsorship are outlined in Table 1. The hypothesis that the rate or response using bank sponsorship was greater than university sponsorship was rejected ( $p < .05$ ). However the hypothesis that the response rate using bank sponsorship was greater than research agency sponsorship could not be rejected ( $p < .01$ ).

TABLE 1

## The Effect of Sponsorship on Mail Survey Response Rates

<u>Treatment Group</u>	<u>Number of Questionnaires Returned</u>	<u>Number of Questionnaires Not Returned</u>	<u>Response Rate</u>
A (bank)	79	93	45.9
B (university)	65	98	38.7
C (agency)	49	112	29.2

## Group Comparability

The differential response rates to the sponsorship treatments could cloud the issue of whether sponsorship affects response bias. An important question is: does sponsorship affect the quality of the responses or just the composition of the self-selected groups who return the questionnaire? While the response rates shown in Table 1 are generated by using a field experiment, the subsequent comparison of the quality of the results utilizes a quasi-experimental design. The results of the quality comparison envisaged by this research are still of great interest to the consumer researcher as bias in results, regardless of source, is relevant.

The three groups comprised of respondents who returned the questionnaire were compared demographically. After appropriate transformation of the five demographic variables (age, sex, children, income, education) stepwise discriminant analysis was conducted.

The groups were not found to be significantly different on these demographics. This finding gives greater confidence to the view that differential response rates through self-selection did not affect the group composition.

## Response Bias

The question of bias was examined using stepwise discriminant analysis comparing both bank and university responses and bank and research agency responses. The independent variables were thirteen items evaluating the bank's performance in the past year.

Only two of the thirteen items were significantly different ( $p < .05$ ) across the bank and university groups and are shown in Table 2. A separate one-way analysis of variance was conducted to test for equality of group means on each of the thirteen evaluation variables. None of the variables were different across the two groups. The university group considered the bank to have less friendly new accounts staff but better overall service.

Multiple discriminant analysis was also used to differentiate the bank group and the research agency group on the thirteen evaluation items. Results, outlined in Table 3, indicate that only one of the variables was significantly different ( $p < .05$ ) across the groups. Separate one-way analyses of variance confirmed this result. The agency group evaluated the timeliness of service higher than the bank group.

TABLE 2

## Evaluation Variables Which Were Significantly Different Between Bank and University Groups

<u>Variable</u>	<u>Bank Mean</u>	<u>University Mean</u>	<u>Partial F</u>
Friendliness of new accounts staff	4.78 <sup>a</sup>	4.69	6.09
Overall customer service	4.45	4.81	9.55

<sup>a</sup>where 1 = very poor, 6 = excellent

TABLE 3

## Evaluation Variable Which Was Significantly Different Between Bank and Research Agency Groups

<u>Variable</u>	<u>Bank Mean</u>	<u>Agency Mean</u>	<u>Partial F</u>
Time waiting in line for service	3.49	3.97	4.51

## Discussion and Conclusions

It was found that sponsorship does affect response rate. Commercial (bank) response rate was significantly higher than research agency response rate but not higher than university agency response rate from random samples of commercial customers.

After establishing that no differences existed between these resultant groups (those who returned the questionnaire from each treatment group) on selected demographic variables, the quality of the resultant responses was compared. Although the analysis did show a statistically significant difference between the groups on a few variables, the multivariate analysis results did not indicate a consistent sponsorship bias. On the variable, "Friendliness of new accounts staff," the mean evaluations by bank sponsored respondents were higher than those by university sponsored respondents while on the second variable, "Overall customer service," the mean evaluation by the university group was higher than that by the bank group. There were no differences for the 11 remaining evaluation scales, e.g. friendliness of tellers, speed and efficiency of tellers, accuracy, drive-in facilities, willingness to make loans. Univariate analysis found no differences on any of these items. When bank and research agency responses to bank evaluation items were compared only one variable was different. Mean evaluation on the variable "Time waiting in line for service" was higher for the research agency sample than the bank sample. Here customers aware of the direct link between the sponsor and their own consumer behavior were more critical of the bank. However no differences on the remaining twelve evaluation items were found using either univariate or multivariate analyses. It appears that evaluations gathered from research which is sponsored by a commercial firm are not upwardly biased when compared to evaluations gathered from a university (a highly credible, public sector sponsor) or from an independent research firm (where the true sponsorship of the research was also disguised). This study indicates that in this research setting, commercial sponsorship has a favorable effect on the quantity of the response but no practical effect on the quality of the response.

## References

- Baur, E. Jackson, "Response Bias in a Mail Survey," Public Opinion Quarterly, 11 (Winter, 1947), 594-600.
- Brunner, Allen G. and Carroll, Stephen J., Jr., "The Effect of Prior Notification on the Refusal Rate in Fixed Address Surveys," Journal of Advertising Research, 9 (March, 1969), 42-44.
- Houston, Michael J. and Ford, Neil M., "Broadening the Scope of Methodological Research on Mail Surveys," Journal of Marketing Research, 13 (November, 1976), 397-402.
- Kanu, Leslie and Berenson, Conrad, "Mail Surveys and Response Rates: A Literature Review," Journal of Marketing Research, 12 (November, 1975), 440-453.
- Scott, Christopher, "Research of Mail Surveys," Journal of the Royal Statistical Society, 124 (1961), Series A, Part 2, 143-191.



Gordon G. Bechtel, University of Florida

## Introduction

The papers in the present session all address biases in attribute ratings of attitudinal objects. However, each set of investigators has chosen different object domains and biases for study. Best, Hawkins and Albaum look at ratings of university institutions as a function of the number of scale categories. Miller and Turner study the effects of perceived survey sponsorship upon the attribute ratings of banking service. Finally, Horowitz and Golob investigate shifts upon attribute scales attributable to unreliable survey respondents. Their attitudinal objects are transportation vehicles, both of the kind we know and futuristic. After a consideration of each paper in turn, an attempt will be made to gain some broader perspectives, as well as directions for future research upon categorical survey responses.

The Effect of Varying Response Intervals  
on the Stability of Factor Solutions  
of Rating Scale Data

Here the treatment conditions consist of two scale formats, one with five categories, and the other a continuous graphic scale (theoretically with an infinite number of categories). The authors study two types of effects, univariate and multivariate, which represent important issues in contemporary survey measurement.

## Univariate Effects

Under each response format the treatment group rated three objects (e.g., the university bookstore) upon ten attributes (e.g., dependable). Figure 1 plots the 30 item means under the discrete format upon the corresponding means observed under the continuous format. Although the form is linear and the correlation is .92, we should note that this univariate equivalence is still rather modest, especially since these rating scale means represent aggregate measurement. For example, Jones (1960) used the Method of Successive Intervals in large-scale consumer surveys involving 20 food items. Two rating forms were employed, one containing nine categories and the other six, with labels representing varying degrees from "dislike" to "like". Each sample size was approximately 900, and the plot of the six category scale upon the nine category scale was almost a perfect straight line. Of course, we are left to wonder whether this result is due to Jones' superior sample size or to the superiority of the Method of Successive Intervals as an alternative to the more usual use of rating scale means.

## Multivariate Effects

Whatever the degree of scale invariance in Figure 1, the authors' multivariate results appear to indicate much larger treatment effects. For example, the (varimax-rotated) factor loadings for one attitudinal object, the university, are plotted for the two treatment conditions in Figure 2. Here we see nothing like the correspondence observed for the univariate results in Figure 1.

However, the use of rotated factor loadings here could be clouding the issue of multivariate correspondence. Principal-components loadings would provide a safer comparison, due to their uniqueness in representing a correlation matrix; better yet, the most direct comparison possible would be given by the plot one correlation mat-

rix upon the other, i.e., the plot of the  $\binom{10}{2} = 45$  correlations for the discrete rating condition upon the corresponding correlations observed for the continuous, graphic ratings. This type of direct comparison has been used by Andrews and Withey (1976) to assess the invariance of multivariate relationships in social-indicators research.

If these further probes should uphold the low level of association in Figure 2, then it would appear that the authors have put their finger on a particular univariate-multivariate differential in rating scale work. It is interesting to note that other multivariate effects in surveys have been cited by Turner and Krauss (1978), who indicate that "context and wording artifacts may disrupt the pattern of correlations between subjective indicators and other variables..." Of course, context and wording represent stimulus effects, while the number of rating categories studied here is on the response side.

The Effects of Sponsorship on Mail Survey  
Response and Evaluation Bias

In the Miller and Turner paper we move over to the stimulus side in considering the contextual effect of perceived questionnaire sponsorship. This is a field experiment at the outset with three "sponsorship" groups being randomly drawn from the checking account file of a particular bank. Under each condition of perceived sponsorship the respondent rated 13 attributes of bank service, which is the only attitudinal object considered. However, prior to the analysis of the ratings, the authors evaluate the treatment effects upon response rate itself. These two types of dependent variables lend a very interesting quality to this study in that the first defines a true field experiment, while the second derives from a quasi-experiment carried out on the returned questionnaires in each treatment group.

## The Field Experiment upon Response Rate

Although all three treatment groups had response rates below 50%, a significantly higher rate of return was observed for the group receiving a cover letter signed by a bank officer. The authors' conclusion that "commercial sponsorship has a favorable effect on the quantity of response" would seem to be too strong, however, on the basis of these results. When the sponsor is your own bank, asking questions about banking service, the higher response rate is not surprising, and it may not generalize to other survey settings. For example, it is an open question as to whether or not a General Mills food-preference survey would garner a higher response rate than an identical survey carried out by the USDA.

## The Quasi-Experiment with Attribute Ratings

In the subsequent comparisons of the attribute ratings, the three treatment groups are no longer randomly equivalent. However, despite the less than 50% return in each group, a discriminant analytic check revealed that the three sub-groups of returnees were still demographically comparable. On this basis the authors proceed to Table 2, which reports the results of two stepwise discriminant analyses, each involving a comparison of two treatment groups. In these analyses only three of the 13 satisfaction items surface as discriminators.

Moreover, two of these three items wash out when the discriminant analyses are checked by means of a separate one-way analysis of variance for each item. Thus it would seem that perceived sponsorship had virtually no contextual effect upon rated attribute satisfaction.

This latter result could have been more compactly shown by substituting (for Table 2) a single, three-group discriminant analysis parallel to that carried out with the demographic variables. This would be attended, of course, by 13 corresponding analyses of variance, each involving three groups. Putting this technicality aside, however, the present application of the stepwise discriminant procedure, with the ANOVA check, represents an interesting approach to the selection of attribute-satisfaction items in survey research. This technique might be used in its own right to select attributes which discriminate between, say, known demographic segments.

#### Survey Data Reliability Effects on Results of Consumer Preference Analyses

Thus far we have looked at the response format itself, as well as stimulus-context effects upon questionnaire responding. The paper by Horowitz and Golob, in contrast, spotlights the survey respondents themselves as potential sources of distortion. Their consideration of respondent unreliability as a measurable "trait" places this personal characteristic among other response dispositions, such as acquiescence and social desirability. Hence, this paper is also a contribution to the literature of personality assessment.

#### The Trait of Unreliability

Unreliability is first approached here through a subsample of 56 individuals (of the panel of 1,565), who were retested with 87 items. A test-retest correlation (over items) was calculated for each individual, and the distribution of these coefficients appears in Figure 1. This distribution is then used to dichotomize the subsample into "reliable" and "unreliable" respondents. Subsequent study of these protocols indicated that the unreliable respondent tends toward (1) a high usage of a particular response category, and (2) a high proportion of unanswered items. Thus, an individual's questionnaire can be scored for these two characteristics, just as it can be scored for other response dispositions. We should note, however, that the first characteristic is not an independently discovered property of unreliable respondents. That is, single category usage depresses intra-individual variation over items, which, in turn, depresses the test-retest coefficient used to classify the individual as unreliable. In any event, since each questionnaire can be scored in this way, the authors are able to identify unreliable respondents within the panel of 1,565 measured at the first time point only.

#### Application of Unreliability Scores

Once so identified, the unreliable respondents were eliminated from the initial panel, and the effects of this deletion upon the (aggregate) scale distributions were observed. First, when radically different, futuristic transportation concepts are rated, this elimination shifts the distribution outward toward the nearest endpoint of the scale. This shift is attended by a decrease in the dispersion of the distribution, which, however, would appear to be an artifact associated with the truncated nature of scale itself.

Next, when presently-owned vehicles are rated, opposite shifts occur. Finally, for an in-between concept, i.e., one which is rather similar to existing automobiles, the elimination of unreliable respondents leaves the distri-

bution essentially unchanged. Thus, the authors demonstrate a systematic interaction between respondent unreliability and the degree of departure of product concepts from existing products. Also, since the response distributions appear, in the main, to be on the positive side of the various scales, unreliable, i.e., older and less educated, respondents, would seem to be more favorable toward existing vehicles and less favorable toward radically different ones.

In a final assessment of the effect of unreliability, Horowitz and Golob relate ownership preferences among transportation concepts to satisfaction with fuel economy and vehicle size, which were found to be "representative" attributes. A multinomial logit model (the details of which are not given) was used to estimate attribute weights, which were essentially unaltered by the elimination of unreliable respondents. Of course, parameter robustness to unreliability must be established for each model and attitudinal domain separately, but the present paper outlines the method for doing this with survey samples. The very availability of this procedure, however, raises the question of what to do if the model is affected by unreliability. That is, which set of parameter estimates should we keep; those for the whole sample, which are more representative, or those for the subsample, which are more reliable? Since this is much broader than a modeling issue, we will return to it below.

#### Future Directions

Looking back upon these three papers, as well as rating-scale studies in general, it is clear that the distinction between aggregate and individual measurement is crucial. In the aggregate case Horowitz and Golob found no differences between test-retest averages, but their cross-sectional correlation for each test-retest pair fell well below .50! This different order of magnitude between aggregate and disaggregate relationships is commonly seen in survey research. For example, in the area of consumer optimism, aggregate time-series regressions display multiple correlations in the .90's, while their cross-sectional counterparts fall into the .40's (Strumpel, Morgan, and Zahn, 1972). In the USDA's consumer satisfaction surveys the aggregate regression of food satisfaction upon attribute satisfaction shows an  $R^2$  of .98. This coefficient falls to .64 in the corresponding cross-sectional regression (Bechtel, 1978).

The message here is clear and well-known, i.e., aggregate measures are more stable over time and also lend to higher degrees of prediction and/or explanation. However, this statistical advantage of aggregate measurement at the societal level does not solve our measurement problems at the individual level. Even so, the numerical rating scale, with its assignment of integers to ordered categories, continues to enjoy almost ubiquitous usage. These numerical assignments involve scale truncation, as well as the (perhaps unwarranted) assumption of equal perceptual steps between successive categories. In a sense, this is measurement by fiat, since there is no way of assessing the fit, i.e., of evaluating, the measurement procedure itself.

It is possible that stochastic response models, with parameters estimated by maximum likelihood methods, can alleviate these problems at the individual level of questionnaire measurement (cf. Cox, 1970). In the aggregate case, however, clear alternatives to numerical rating methods are already available. Here there is a natural grouping of questionnaire responses such that we are able to observe the proportion of respondents falling into each scale category. These proportions circumvent the assignment of integers to categories by enabling us to substitute a probabilistic re-

sponse model for the usual rating method. This model can be parameterized to represent object or attribute values, as well as the reference points upon the rating continuum. Therefore, estimates of object or attribute parameters now replace the mean ratings usually calculated.

A principal advantage of model-based measurement lies in the fact that the measurement procedure itself is vulnerable to a goodness-of-fit test, which assesses the scalability, and therefore the quality, of the data itself. This approach, known in psychometrics as the Method of Successive Intervals, is described by Torgerson (1958), and its application to food-preference surveys by Jones (1960) has been noted above. Also, a generalization of successive-intervals scaling in the form of a logistic response model has recently been presented (Bechtel, in press). This generalization permits us to study contextual effects, such as perceived sponsorship, upon scale values and upon scalability itself under several experimental conditions.

Finally, whether stochastic response models or more standard methods are used for survey measurement, we will continue to have less control over response dispositions than we have over stimulus biases. Obviously, questionnaires can be pilot tested and corrected for stimulus effects such as context or wording, but respondent unreliability is out of our hands. The tactic of throwing out unreliable survey respondents seems very severe, especially since they are older and less well educated. This demographic deletion aggravates the non-response problem, already serious in survey research (ISR Newsletter, 1976), by further warping carefully drawn probability samples. Of course, the problem of sample representativeness involves the ultimate application of present research findings, which is somewhat beyond the scope of this session. The paper by Horowitz and Golob has indeed contributed to our thinking about response error by placing unreliability among other potentially biasing response dispositions, such as acquiescence and social desirability. The continuous nature of all of these traits, though, points up the further problem of establishing arbitrary cut-offs in purifying survey data. Whether the benefit of this purification will outweigh the cost in representativeness is an important matter for future inquiry.

#### References

Frank M. Andrews and Stephen B. Withey, Social Indicators of Well-Being (New York: Plenum Press, 1976).

Gordon G. Bechtel, "Consumer Satisfaction with Foods and Their Attributes," Working Paper No. 6, Center for Consumer Research, College of Business Administration, University of Florida, 1978.

\_\_\_\_\_, "A Scaling Model for Survey Monitoring," Evaluation Quarterly, in press.

D. R. Cox, Analysis of Binary Data (London: Methuen, 1970).

ISR Newsletter, "The Growing Interest in Telephone Interviewing," Institute for Social Research, The University of Michigan, 4 (Autumn, 1976), 2.

Lyle V. Jones, "Some Invariant Findings under the Method of Successive Intervals," In H. Gulliksen and S. Messick (eds.), Psychological Scaling: Theory and Applications (New York: John Wiley and Sons, 1960).

B. Strumpel, J. N. Morgan, and E. Zahn (eds.), Human Behavior in Economic Affairs (Amsterdam: Elsevier, 1972).

Warren S. Torgerson, Theory and Methods of Scaling (New York: John Wiley and Sons, 1958).

Charles F. Turner and Elissa Krauss, "Fallible Indicators of the Subjective State of the Nation," American Psychologist, 33 (May, 1978), 456-470.

Paul E. Green, University of Pennsylvania

## Introduction

One of the common threads of the three papers comprising this session is that they all employ rating scales as a major aspect of the study. The historically much maligned rating scale has managed not only to survive its many critics but, indeed, to flourish and proliferate with the intensity of unbottled fruit flies. Such resilience and ubiquity should not go unrewarded, as these papers attest.

The Effect of Varying Response Intervals  
on the Stability of Factor Solutions  
of Rating Scale Data

The Best, Hawkins, and Albaum paper examines the relationship between two types of rating scale intervals--a five-point scale and a continuous scale--from both a univariate and multivariate standpoint. The univariate results are fully in accord with expectations. The authors find very high correspondence between scale means; the Kolmogorov-Smirnov test for correspondence of distributions reinforces this finding.

In contrast, the multivariate portion of the study--entailing separate factor analyses of the two types of rating scales and subsequent correlations of factor loadings between scale types--showed poor results. However, before concluding that factor analysis is all that sensitive to the numbers-of-intervals characteristic, a few comments might be raised regarding the authors' methodology:

1. Assuming that the authors applied principal components analysis to the correlation matrix (followed by Varimax rotation), why did they not correlate the (off-diagonal) correlations between the two types of scales? This would provide a direct summary of how closely related the two types of scales were and, of course, could be done for each of the three concepts separately.
2. The factor loadings summaries of Table 2 do not show particularly clean patterns for each scale type separately. One wonders if the "right" number of factors were extracted in the first place. Did the authors check the between-scale correspondences for, say, the 2-factor and 4-factor solutions?
3. Perhaps the analysis summarized in Table 2 should have been augmented by a factor matching procedure. Assuming that the authors would wish to retain orthogonality of factors, they could use either Cliff's procedure (Cliff, 1966) or the closely related Schönemann and Carroll factor matching approach (Schönemann and Carroll, 1970).

In summary, there are some unsettled questions about the lack of congruence between the two types of ratings scales that might be related to the factor analyses themselves. If the preceding suggestions are implemented and the results still support the authors' earlier conclusions, then additional study of the problem is clearly indicated. Perhaps this additional study could include Monte Carlo simulations as well as the type of empirical comparisons reported here. Finally, from a behavioral viewpoint, we might ask: why would scale interval type be expected to affect covariances, but not the means or variances?

The Effects of Sponsorship on Mail Survey  
Response and Evaluation Bias

The paper by Miller and Turner also employs rating scales as a central feature of the analysis. The authors have turned out an attractive field experiment dealing with the question of whether type of sponsorship affects mail survey response rate, the demographic composition of those who do respond, and rating scale response patterns.

For the most part the study was nicely designed and executed. However, a few critical comments--hardly more than nitpicks--might be offered as a space filling effort:

1. Were such background variables as number of years with current bank, number of different bank services used, or average size of checking account balance included in the survey? If so, perhaps this type of variable might show greater sensitivity to survey sponsorship than the more traditional demographics.
2. In the Table 1 analysis, one might wish to conduct a oneway ANOVA, followed by multiple comparison tests (assuming that the overall ANOVA results are significant).
3. Why did the authors not run a three-group discriminant analysis--similar to what was done in the case of the demographics--in examining the significance of mean differences across rating scales?
4. One wonders if the Table 2 findings are useful, given no significant univariate F values. At the very least, additional discriminant runs should be made to see if the two variables are significant together and, if so, whether both are needed. Model comparison tests could be used for this purpose (Rao, 1952).

However, these are quite minor suggestions that do not affect the substantive conclusions of the paper. The study deals with a set of interesting issues in survey research. One naturally wonders how well the results on response bias (actually, the lack thereof) generalize to other contexts and respondent populations.

Survey Data Reliability Effects on Results  
of Consumer Preference Analyses

The paper by Horowitz and Golob addresses an important problem in the use of rating scales, namely their test retest reliability. The authors' study is quite competently carried out and my critical comments are few in number and minor in import. However, some points come to mind which the authors might wish to consider:

1. How did the authors arrive at a sample of 56 respondents for test/retest purposes? Perhaps two independent samples could have been drawn from purposes of assignment rule validation. If this were done, the assignment rule portrayed in Table 1 could have been checked against fresh data before going to the main sample.
2. What appears to be missing is information on which of the 29 scales show the greatest incidence of missing data, which sets of scales show the highest (relative) incidence of unvarying response categories, or whether the concepts themselves differ in terms of the incidence of these putative sources of

unreliability. In other words, while we may know that missing data and low variance in response category selection are associated with low reliability, how can we design respondent tasks to reduce the tendency for these bad things to occur?

3. When one returns to the total sample, it would also seem appropriate to split that sample into halves and carry out some type of cross validation of the new assignment rules adopted in Table 2.
4. Incidentally, what is the rationale for the four types of rules in Table 2? Also, what is the precise nature of the significance tests being carried out? Do they involve the subsample mean versus the rest of the sample (for each subsample mean, in turn), or what?
5. Insofar as Hypothesis I is concerned, how big are the effects? Clearly, with a sample size this large the results could be statistically significant but not operationally important.
6. Hypothesis II makes me a bit uncomfortable. The study shows that unreliability and age/education are associated. But how about reliable older, poorly educated respondents? Do they also show the same stereotype evinced by unreliable older, poorly educated respondents?
7. As a matter of interest, why all the concern about Hypotheses I, II, and III in the first place? It seems to me that if reliability (as defined in this study) is really important, it is not appropriate nor useful to average over the responses of reliable and unreliable respondents. A more sensible thing, of course, is to analyze both sets of data separately. I assume that this is what most researchers do currently when they have reason to suspect different reliability levels across respondents and assignment rules are available for classifying respondents according to their degree of reliability. (Degree of reliability may simply reduce to dichotomous assignments in some cases.)
8. Unfortunately, it is not clear to me how the multinomial logit model was applied in this example; perhaps the authors could spend some time in showing us how they developed the predictor-variable set and how the criterion variable was defined. However, it is not surprising that the type of unreliability found in this study did not affect the logit model results.
9. On a more general basis it has been known for a long time that random error added to regression-like clinical judgment models does not affect the relative sizes of the partial regression coefficients. Similarly, in MDS programs, such as Carroll and Chang's INDSCAL, the addition of random subjects has little effect on the group stimulus space, even though the overall fit is (obviously) reduced.

All in all, this is an interesting study that deserves follow up by other researchers. In particular, I would like to see what other clues to unreliability might be found in other survey research and what background correlates (comparable to age and education in the current study) are noted.

#### Future Research Possibilities

When all three papers are considered together, what lessons can be learned? Bearing in mind that the papers were prepared independently and the session theme is a broad one, it still seems to me that a few suggestions

for further research can be made.

For example, both the Best, Hawkins, and Albaum study and the Horowitz and Golob research might benefit from comparison studies that utilize Monte Carlo simulation. Clearly, the Monte Carlo technique could be used to examine questions related to number of scale intervals and other kinds of scale transformations on factor loadings matrices and on "true" versus spurious factors. Moreover, Horowitz and Golob could use Monte Carlo methods to see what kinds of unreliability affect the parameter estimates derived from such decomposition models as the multinomial logit, and how the estimates are affected.

A second area, of potential relevance to all three papers, involves the design of appropriate rating scales and perceptual/evaluative tasks in the first place. Are people more comfortable using one type of rating scale than another? Can various devices be used to reduce the incidence of missing data and same-category checking? Should rating scale design vary with the type of sponsor identified in the study? Of course, the list is virtually endless since so few general principles have been adduced to date, despite the fact that rating scales are almost de rigeur in any self-respecting survey that goes out these days.

Not to end on an unduly heretical note, but perhaps respondents are being asked to rate just too many things on too many scales, where the scales are largely redundant to begin with or are too vague and nonoperational to be useful, once the results are in. Certainly, those working in psychographics have had cause for concern regarding the ability of clusters based on life style rating patterns to predict other aspects of the consumer, such as brand choice. More importantly, can attribute rating scales--no matter how cleverly designed and executed--provide rich enough response data to enable us to design an "optimal" university book store, a consumer banking service, or a new automobile? My own experience with attribute ratings suggests that this approach to product/service design leaves much to be desired.

Clearly, rating scales for operationalizing perceptions of choice objects have a long history in expectancy-value modeling and the like, and it is not likely that this approach will soon be replaced by conjoint techniques (Green and DeSarbo, 1978), or other such contenders. Still, one wonders if other ways of eliciting perceptions can be found that provide a sounder basis for product/service design and better predictions of preferences for new objects. At least it might be fun to look.

#### References

- Norman Cliff, "Orthogonal Rotation to Congruence," Psychometrika, 31(1966), 33-42.
- Paul E. Green and Wayne S. DeSarbo, "Additive Decomposition of Perceptions Data via Conjoint Analysis," Journal of Consumer Research, 5(June, 1978), 58-65.
- C. R. Rao, Advanced Statistical Methods in Biometric Research (New York: John Wiley and Sons, 1952).
- Peter H. Schönemann and Robert M. Carroll, "Fitting One Matrix to Another Under Choice of a Central Dilation and a Rigid Motion," Psychometrika, 35(1970), 245-256.

ESTIMATING THE EFFECTS OF ADVERTISING:  
APPLICATION TO A SOCIAL PROBLEM

James G. Barnes, Memorial University of Newfoundland  
Jacques C. Bourgeois, Carleton University

Abstract

"Per capita" consumption of alcohol has increased dramatically over the last twenty years; for instance, between 1954 and 1974, the total intake of raw alcohol has more than tripled. To many, this social problem is cause for concern. This paper investigates those variables which may have had a significant impact on this increase. It was found that non-controllable variables (e.g., income, unemployment, and urban/rural mix) were generally more significantly related to higher per capita consumption than were marketing/controllable variables (e.g., advertising, distribution, price).

The Problem

Recent years have seen a marked increase in the per capita consumption of beverage alcohol products in Canada. In the year ending March 31, 1974, for example, the average Canadian aged 15 and above consumed 25.24 Imperial gallons of beer, 1.67 gallons of wine, and 2.11 gallons of spirits. Since 1954, the annual per capita intake of raw alcohol (from beer, wine and spirits) has increased by more than 57 per cent (see Table 1).

TABLE 1

PER CAPITA CONSUMPTION OF BEVERAGE ALCOHOL PRODUCTS IN CANADA, BY PERSONS AGED 15 AND ABOVE: 1954, 1964 and 1974

	1954	1964	1974	% Increase 1954 to 1974
Per Capita Consumption (in gallons) of:				
BEER	19.61	20.68	25.24	29.5
WINE	0.52	0.79	1.67	221.2
SPIRITS	1.06	1.36	2.11	99.1
Per Capita Intake of Raw Alcohol (in gallons) from:*				
BEER	0.98	1.03	1.26	
WINE	0.08	0.13	0.27	
SPIRITS	0.42	0.54	0.84	
Total Per Capita Intake of Raw Alcohol (in gallons)	1.48	1.70	2.37	60.1

\*The approximate percentage of raw or absolute alcohol in beer is 5%, in wine 16% and in spirits 40%.

SOURCE: Statistics Canada, The Control and Sale of Alcoholic Beverages in Canada, Cat. No. 63-202, Annual.

Such statistics have been considered cause for alarm by many groups. The social costs associated with the rapid increase in consumption of beverage alcohol products have been widely debated. Lalonde (1975) estimated that seven per cent of the drinking population consume 40 per cent of all alcohol sold. This translates into 15 gallons of raw alcohol consumed each year by each person in the heavy drinking population. It is the alcoholic and the occasional heavy drinker who contribute most to the onerous social costs associated with the consumption of beverage alcohol products.

The Commission of Inquiry into the Non-Medical Use of Drugs concluded in their Final Report (1973) that

alcohol is, and is likely to remain, Canada's most serious non-medical drug use problem ... From almost any point of view the effects of the excessive use of alcohol are more harmful than those of any other form of non-medical drug use. (p. 37)

The problems associated with excessive use of alcoholic beverages have been discussed under three major headings: physical health problems, mental health problems, and social problems.

It has been clearly demonstrated that excessive use of alcohol contributes to diseases such as cirrhosis of the liver, hepatitis, cancer of the upper respiratory and digestive tracts, tuberculosis, and malnutrition. Heavy use of alcohol has similarly been associated with a variety of psychiatric and neurological disorders. In Canada in 1972, 17 per cent of first admissions and 16 per cent of readmissions to psychiatric wards and hospitals were for alcoholism and alcoholic psychosis.

The social problems associated with excessive alcohol use are numerous. The Commission of Inquiry into the Non-Medical Use of Drugs (1973) concluded that, of all drugs used medically and non-medically, alcohol has the strongest and most consistent relationship to crime. In their Final Report (1973) this Commission cited a recent Philadelphia study which indicated that alcohol was present in either the offender or the victim in 64 per cent of homicide cases over a five-year period. Similarly, the contribution of alcohol use to motor vehicle accidents has been well documented. About half of all drivers killed in automobile accidents have been found to have positive blood alcohol levels.

In industry, excessive alcohol use contributes to high rates of absenteeism and to numerous on-the-job accidents. Equally serious, but considerably more difficult to quantify, are the problems which are imposed by problem drinkers on their family and friends. The impact of the alcoholic parent on the lives of his or her children may be significant.

Some attempts have been made to quantify the social costs associated with excessive use of alcohol. The costs of lost production, health and medical care, motor vehicle accidents, alcohol programs and research, criminal justice and social welfare systems have been estimated at over \$25 billion per year in the United States and \$2.5 billion per year in Canada (Health and Welfare Canada, 1975). Since these estimates were made in 1971, it is likely that the total identifiable costs of

alcohol-related problems today are closer to \$40 billion in the United States and \$4 billion in Canada.

#### Advertising's Contribution to Increased Consumption

The question of the effect of advertising on aggregate demand has been debated at length. In the tobacco and alcohol industries in particular, critics have been advocating for some time that advertising should be reduced or eliminated. The assumption obviously is that such a reduction in advertising would lead to a corresponding reduction in aggregate demand or per capita consumption. But studies of the impact of advertising on total sales of cigarettes have tended to indicate that advertising's effect on industry sales can not be demonstrated. Schmalensee (1972) concluded that "advertising seems unlikely to be a very powerful force in the cigarette industry". In Canada, per capita consumption of cigarettes has continued to increase, despite the fact that manufacturers voluntarily withdrew all cigarette advertising from the broadcast media in the late 1960's.

The lack of proof concerning the effect of advertising on industry sales has not stopped the critics of advertising from proposing that the advertising of all alcoholic beverages be banned. The association of advertising with increasing consumption of alcoholic beverages has not been made only by religious groups or social activists. Some leading politicians have been outspoken in their criticism, not only of the sheer volume of advertising, but also of the fact that much advertising, especially for beer, presents the product in the context of a glamorous life style, suggesting that use of the product is essential to having a good time (Lalonde, 1974).

The criticism of beverage alcohol advertising is very often directed toward the observation that advertising leads young people to start drinking. The very rapid increase in consumption of alcoholic beverages among young people in recent years is often cited as an indication of the influence of advertising (Ontario Youth Secretariate, 1976). A more plausible explanation of the increase might be found in the fact that all ten Canadian provinces have reduced the legal drinking age from 21 to 18 or 19 within the past five years, although two provinces have recently raised it again from 18 to 19.

Advertising of alcoholic beverages is regulated in Canada by both the federal and provincial governments. At present, advertising of spirits is not permitted on radio or television and all beer advertising in these media is subject to the approval of the Canadian Radio-Television and Telecommunications Commission. Media other than broadcast are regulated by provincial governments and the stringency of such regulations is quite variable. For example, three provinces in effect prohibit all forms of advertising for beverage alcohol products, while one province has no regulations dealing with such advertising.

#### Variables Influencing Per Capita Consumption

The social problems associated with the excessive use of alcohol are also associated primarily with the alcoholic and the heavy drinker. A number of authors have explored the possibility that the distribution of alcohol consumption is positively skewed with a curve approximating the lognormal distribution (Miller and Agnew, 1974). If such a distribution applies, then the number of heavy drinkers and alcoholics in a population may be determined as a function of the mean (per capita) level of consumption of alcohol. This observation has led certain authors to suggest that the best way to reduce the number of heavy drinkers, and the problems associated with excessive use of alcohol, is to reduce the mean

level of consumption in the population (Schmidt and de Lint, 1971). Before a program can be launched to reduce per capita alcohol consumption, some information must exist concerning those factors which influence the mean level of consumption.

Governments in many jurisdictions are currently being pressured to take some action to reverse the trend toward increased per capita consumption of alcoholic beverages. But, before governments can take any action, it would be appropriate for them to attempt to determine which factors are most important in influencing per capita consumption. The factors which have been associated with per capita consumption of beverage alcohol products may be divided into two categories: (1) the controllable or 'marketing' variables such as price, advertising, and availability, which may be influenced by governments in the short run; and (2) the uncontrollable or socio-economic variables such as per capita income, urbanization rate, and ethnic mix of population, which are not subject to change in the short run.

#### Availability

Various aspects of distribution or availability have been linked to levels of per capita consumption. Schmidt (1972) suggested that recent increases in consumption are a result of reductions in the legal drinking age. Popham, et al. (1971) have concluded that per capita consumption of alcoholic beverages is unaffected by the hours of sale in liquor stores and by the number of outlets at which liquor is sold. Simon (1966b) concluded that per capita consumption is not related to the fact that alcoholic beverages are sold by state monopoly rather than by private enterprise. Smart (1974) and Goldberg and Gorn (1975) have observed different purchase patterns by customers in self-serve liquor stores as compared with customers in conventional counter-service stores.

#### Price

Much has been written about the impact on the level of per capita consumption of alcoholic beverages. Seeley (1960) has shown that death rates from cirrhosis of the liver have increased as the price of alcohol (expressed as a percentage of personal disposable income) declined. Since the level of government taxation significantly influences the price of alcoholic beverages, a number of attempts have been made to measure the price elasticity of demand for alcoholic beverages (Simon, 1966a; Johnson, 1973). The results of these attempts have been largely inconclusive, as they tend to vary with the country studied, size of sample, and the objectives of the study. Several authors (Popham, et al., 1971; Johnson, 1973) have concluded that the tax structure can be used as a valuable tool to reduce per capita consumption of alcoholic beverages. The price of alcoholic beverages relative to personal disposable income has been declining for at least forty years and the per capita consumption of these products and the levels of alcoholism in the population continue to rise.

#### Advertising

Despite the ongoing criticism of beverage alcohol advertising and the demands for increased regulation, few studies have been devoted to examining the question of advertising's effect. Globetti (1973) discussed the potential of alcohol education programs as measures to counteract advertising for alcoholic beverages. Other authors have examined the effect of cigarette and proprietary drug advertising on per capita consumption of cigarettes and on drug abuse. Hulbert (1974) concluded that television and the advertising of proprietary drugs on television do not appear to be important influences on usage of illicit drugs.

## Socio-economic Variables

A series of socio-economic variables has also been associated with per capita consumption of beverage alcohol products. Wechsler, et al. (1970) examined the drinking behavior of various religious and ethnic subgroups and found that alcohol consumption was significantly related to religious-ethnic subgroup membership. Bales (1946) observed that factors affecting problem drinking are often culturally based. Cisin and Cahalan (1971) concluded that whether or not a person drinks is primarily related to socio-cultural rather than psychological variables. They concluded that socio-economic status and urbanization are important correlates of heavy drinking. In a recent study, Johnson and Oksanen (1974) incorporated ethnic, religious, educational and occupational variables into a study of per capita consumption and concluded that "these variables, along with traditional economic ones, are significant determinants of the demand for alcoholic beverages" (p. 12). In one of the most detailed studies to date, Cahalan, et al. (1969) associated per capita consumption of alcoholic beverages with education levels, degree of urbanization, and religion, among other variables. He found that the proportion of heavy drinkers was found to be lowest in the income group below \$4,000. per year, and highest in the \$10,000. to \$14,999. range. Heavy drinkers tended to be more common among urban business and semiprofessional people; among men who had completed high school and not graduated from college; and in large cities. Catholics were above average in the proportion of drinkers considered 'heavy drinkers', but the highest proportion of heavy drinkers was among Irish Protestants.

## The Data

Data were obtained primarily from Statistics Canada reports and from the Annual Reports and price lists of the various provincial liquor boards and commissions, with the assistance of the Canadian Association of Provincial Liquor Commissioners. The data were gathered from all ten provinces of Canada for the period 1966 to 1973.

## Dependent Variables

Measures were obtained of per capita consumption of beer, wine, and spirits by province for each of the eight years from 1966 to 1973. A measure of per capita consumption of actual beverage alcohol was also obtained.

- (1) per capita consumption of beer: the measure of per capita consumption of beer represents consumption by residents of each province who are aged 15 or older. It was felt that this was the age group which is most likely to be involved in the consumption of beverage alcohol products. The total provincial consumption of beer and ale for each year was simply divided by the population aged 15 and older to obtain the relevant figure.
- (2) per capita consumption of wine: this figure was obtained in exactly the same way as was the beer consumption figure discussed above. The total provincial consumption of wine was divided by the population aged 15 and older to obtain the per capita consumption rate.
- (3) per capita consumption of spirits: this figure was also obtained by dividing the total provincial consumption of spirits for each year by the provincial population aged 15 and older.
- (4) per capita consumption of beverage alcohol: this is a measure of per capita raw alcohol consumption and is included in the analysis in order to provide a measure of actual alcohol consumption. This is necessitated by the fact that the three major types of beverage alcohol products differ in actual alcohol content and an analysis based simply on gallonage of each product type consumed could be misleading. For example, a person who consumes 15 gallons of beer is obviously consuming less alcohol than is a person who consumes 15 gallons of spirits,

simply because beer is much lower in alcohol content than is spirits.

This measure of actual beverage alcohol consumption is obtained by converting actual volume of beer, wine, and spirits consumed into beverage alcohol volume through application of the following standard conversion factors:

- beer: 5% alcohol by volume
- wine: 16% alcohol by volume
- spirits: 40% alcohol by volume

Thus a person who consumes 20 gallons of beer in a year is actually consuming one gallon of beverage alcohol. Similarly, 10 gallons of wine is equivalent to 1.6 gallons of beverage alcohol, and 15 gallons of spirits is equivalent to 6 gallons of beverage alcohol.

## Independent Variables

Measures were obtained on a total of nine different variables which were hypothesized to be related to per capita consumption of beverage alcohol products.

- (1) price of beer: this variable represents simply the price set by the provincial liquor boards at which a case of 24 bottles of beer may be sold at retail.
- (2) price of wine: this variable is an index of wine prices. A total of twenty popular brands of imported and domestic wines which are sold in all ten provinces were selected and the prices at which these brands were sold in each province for each year were obtained. This variable represents the average price per bottle of these wines.
- (3) price of spirits: this variable was calculated in the same manner as was the wine price discussed above. Twenty popular brands of spirits were selected and their prices in each province calculated. The variable represents the average price per bottle of these spirits.
- (4) percentage rural population: this variable represents the percentage of the populations of each province which is classified as rural dwellers by Statistics Canada.
- (5) per capita income: this variable represents the per capita income of the residents of each province aged 15 and older and is obtained by dividing the gross provincial income by the total population aged 15 and older.
- (6) unemployment rate: this variable represents the average annual unemployment rate for each province as published annually by Statistics Canada.
- (7) number of liquor stores per 100,000 population: this variable simply represents the number of liquor stores operated in each province by the provincial liquor commission, divided by the population aged 15 and older.
- (8) licensed establishments per 100,000 population: this variable represents the number of establishments licensed by the provincial liquor commission to sell beverage alcohol products in the province, divided by the population aged 15 and older.
- (9) advertising regulation index: this variable represents the extent to which each province regulates the advertising and promotion of beverage alcohol products. A province which permits all forms of advertising and promotion and has no regulations or legislation pertaining to the promotion of beverage alcohol products would score 1 on this variable. Conversely, a province which prohibits all forms of advertising and promotion for such products would score 9 on this regulation index. Scores were assigned to each of the ten provinces for each of the eight years covered by this study. The scores were assigned using a 'panel of experts' approach which involved a detailed study of the regulations and legislation of each provincial government and provincial liquor commission for each of the eight years pertaining to any form of advertising or promotion of beverage alcohol products. Detailed descriptions of the regulatory picture in each province, with changes in regulations and laws being noted as they occurred during the eight-year period, were prepared and a group of university professors who were familiar with marketing



and its regulation were asked to rate each province each year on a scale from 1 to 9, depending on the stringency of the provincial regulations. These ratings were then averaged and became the scores for each province on this variable.

Scores on each of the independent variables were obtained for each of the ten provinces for each of the eight years from 1966 to 1973 with a small number of exceptions. Annual reports and price lists were unavailable from four provincial liquor commissions and consequently the 'price of wine', 'price of spirits', 'licensed establishments per 100,000 population', and 'liquor stores per 100,000 population' variables could not be calculated for these provinces for certain of the years in the eight-year period. The analysis of these variables was carried out using complete data only.

The data on the four dependent and nine independent variables from the ten provinces for each of the eight years from 1966 to 1973 were then analysed using a step-wise multiple regression analysis. In all, four different regression analyses were performed, with each of the four per capita consumption measures discussed above as dependent variables.

#### Analysis

##### Per Capita Consumption - Beer

A total of only four independent variables were significantly (at the .05 level) related to per capita consumption of beer. The results of this first regression analysis are presented in Table 2.

TABLE 2

SUMMARY OF REGRESSION ANALYSIS FOR PER CAPITA CONSUMPTION OF BEER

Variables	Regression Coefficient	Beta Coefficient	"t" Statistic
Rural Population Rate	-17.03	-0.72	8.03
Per Capita Income	0.70	0.27	3.17
Unemployment Rate	-0.78	-0.44	4.17
Price of Spirits	1.89	0.30	2.69
Constant	15.23		
$R^2 = 0.813$			
$F_{35}^4 = 37.92; \alpha < 0.001$			

These results indicate that the per capita consumption of beer in Canada's ten provinces is influenced most significantly by the distribution of population in urban centres, by per capita income levels in the provinces, by the level of unemployment, and by the price being charged for competing products, namely spirits. No other variable was found to have a significant effect on per capita consumption of beer.

The regression equation presented in Table 2 indicates that the more urbanized the population of a province, the higher will be the consumption rate of beer among the residents of that province. Those provinces which have a large percentage of their residents considered rural dwellers generally have lower per capita consumption rates for beer. Secondly, more beer is consumed by the residents of those provinces where per capita incomes are highest. Residents of low income provinces generally have lower consumption rates for beer.

Thirdly, unemployment rates appear to influence the level of per capita consumption of beer in a negative direction. Those provinces where unemployment is highest generally have lower per capita consumption rates for beer.

Finally, a substitution effect of sorts seems to operate between beer and spirits in that there is a direct link between the price of spirits and per capita consumption of beer. These data would suggest that as the price of spirits is increased, consumers will consume larger quantities of beer.

The remaining five independent variables were not found to be significantly related to per capita consumption rates for beer. That is, consumption of beer appears to be affected neither by the price being charged for beer in the provinces, nor by the number of liquor stores and licensed establishments, nor by the extent to which the provinces regulate the advertising of beverage alcohol products.

##### Per Capita Consumption - Wine

Four independent variables also entered the regression equation with per capita consumption of wine as the dependent variable. These four variables explained 56 per cent of the variance in the data. The results of this regression analysis are presented in Table 3.

These results suggest that per capita consumption of wine in Canada is influenced significantly by income levels in the various provinces, by the rural/urban distribution of population, by the number of liquor stores operated by the provincial liquor authorities (per 100,000 population), and by the price being charged by the provincial authority for beer. None of the other five independent variables was found to be significantly related to per capita consumption of wine.

TABLE 3

SUMMARY OF REGRESSION ANALYSIS FOR PER CAPITA CONSUMPTION OF WINE

Variables	Regression Coefficient	Beta Coefficient	"t" Statistic
Per Capita Income	0.14	0.44	3.71
Rural Population Rate	-1.97	-0.67	4.29
Number of Liquor Stores	0.04	0.39	3.01
Price of Beer	0.21	0.41	2.97
Constant	-0.23		
$R^2 = 0.560$			
$F_{35}^4 = 11.14; \alpha < 0.001$			

The regression equation presented in Table 3 indicates again that the higher the per capita income in a province, the higher will be the volume of wine consumed by the residents of that province.

Secondly, as was the case in the previous analysis (Table 2), the extent to which the population of a province is urbanized significantly affects the level of per capita consumption of wine. The more urbanized the population of a province, the higher will be the per capita consumption rate for wine in that province.

Thirdly, the accessibility of retail outlets for wine would appear also to significantly influence per capita consumption rates. Those provinces which have larger numbers of government-operated liquor stores per 100,000 population also exhibit higher rates of per capita consumption of wine. Such a relationship does not exist between consumption rates for wine and the numbers of licensed establishments in a province as, presumably, most wine is sold through liquor stores for home consumption.

Finally, a substitution effect is again evident in the link between the price of beer and consumption rates for wine. As the price of beer increases, so also does the per capita consumption of wine as consumers would appear to be substituting wine for beer as the price of the latter is increased.

Per capita consumption rates for wine appear to be significantly influenced neither by unemployment rates, nor by the prices being charged for wine and spirits, nor by the extent to which the provinces regulate the advertising of beverage alcohol products.

#### Per Capita Consumption - Spirits

Only two independent variables were found to be significantly (at the .05 level) related to per capita consumption of spirits. These two variables explained 44 per cent of the variance in the data. The results of this analysis are presented in Table 4.

TABLE 4

SUMMARY OF REGRESSION ANALYSIS FOR  
PER CAPITA CONSUMPTION OF SPIRITS

Variables	Regression Coefficient	Beta Coefficient	"t" Statistic
Per Capita Income	0.13	0.45	3.60
Ad Regulation Index	-0.09	-0.42	3.33
Constant	1.91		

$$R^2 = 0.443$$

$$F_{37}^2 = 14.73; \alpha < 0.001$$

These results indicate that per capita consumption rates for spirits are most significantly affected by per capita income levels. Again, those provinces where personal incomes are highest demonstrated the highest per capita consumption rates for spirits, as they had for beer and wine.

Of the four regression analyses performed, this analysis with per capita consumption of spirits as the dependent variable was the only one where the relationship between the regulation of advertising and the dependent variable was significant. It is suggested in this analysis that the more stringent the regulation of advertising of beverage alcohol products in a province, the lower will be the per capita consumption rate for spirits in that province.

The other seven independent variables failed to produce a significant relationship with the dependent variable. Per capita consumption of spirits appears to be significantly affected neither by the prices being charged for beer, wine and spirits, nor by the number of liquor stores and licensed establishments serving the public, nor by unemployment and population urbanization variables.

#### Per Capita Consumption - Alcohol

An attempt was made to determine those factors which influenced overall alcohol consumption among Canadians. Per capita alcohol consumption figures were obtained by converting consumption rates for beer, wine and spirits into equivalent volumes of beverage alcohol, as discussed earlier.

A regression analysis revealed that four variables were significantly related to per capita alcohol consumption. These four variables explained 68 per cent of the variance in the data. The results of this final analysis are presented in Table 5.

TABLE 5

SUMMARY OF REGRESSION ANALYSIS FOR  
PER CAPITA CONSUMPTION OF TOTAL ALCOHOL

Variables	Regression Coefficient	Beta Coefficient	"t" Statistic
Rural Population Rate	-1.43	-0.60	5.74
Per Capita Income	0.09	0.33	2.97
Price of Spirits	0.24	0.39	2.59
Unemployment Rate	-0.04	-0.25	1.78
Constant	0.79		

$$R^2 = 0.676$$

$$F_{35}^4 = 18.27; \alpha < 0.001$$

These results indicate that overall per capita consumption of beverage alcohol among Canadians is influenced most significantly by the urban/rural distribution of the population in specific provinces, by personal income levels, by the prices charged by provincial liquor authorities for spirits, and by levels of unemployment in the provinces.

This final regression equation presented in Table 5 indicates again that the more urbanized the population of a province, the higher will be the overall rate of per capita consumption of beverage alcohol in that province. Similarly, those provinces which have the highest levels of personal income also exhibit the highest rates of beverage alcohol consumption.

An interesting relationship between the price of spirits and overall consumption of beverage alcohol products may be observed in Table 5. These results suggest that as the price charged by provincial liquor authorities for spirits is increased, a corresponding increase will occur in per capita consumption of beverage alcohol. While this relationship is statistically significant, it is not the relationship which would have been hypothesized to exist. Normally, one would expect consumption to decline, or at least increase at a slower rate, as prices increase. These results, obtained with regard to the effect of the price of spirits on per capita consumption of alcohol, would suggest that other variables are operating through the price variable and which are highly correlated with both the price and consumption variables.

Finally, overall per capita alcohol consumption was found to be significantly affected by unemployment rates. Those provinces which exhibit the highest rates of unemployment among their residents generally exhibit the lowest rates of per capita consumption of beverage alcohol.

## Summary and Conclusions

Those variables which have been found to be most significantly related to per capita consumption rates of beverage alcohol products are generally 'societal' variables rather than 'marketing' variables. Consequently, significant relationships were found between per capita consumption of beer, wine, spirits, and beverage alcohol and such variables as unemployment levels, per capita income and urban density of population. All of these variables are not controllable in the short and medium run by provincial liquor authorities or by manufacturers of beverage alcohol products. Generally, those variables which are controllable such as price, number of liquor stores and licensed establishments, and stringency of advertising regulation were found to not significantly affect per capita consumption rates.

Such results would suggest that attempts on the part of governments to reduce, or at least slow the increase in, per capita consumption rates will have little effect if such attempts take the form of price increases, reduction in the number of licenses issued or liquor stores opened, or increased regulation of advertising of beverage alcohol products. Consumer behavior with respect to beverage alcohol products would appear to be not significantly influenced by traditional 'marketing' variables such as price, distribution, and promotion. While such variables may be influential in affecting consumer choice among brands, this study would suggest that they have little effect on overall rates of consumption of beverage alcohol products. Governments should not expect consumption to decrease dramatically if prices are raised or advertising is restricted. If reduced, or more moderate, consumption of beverage alcohol products is desired, then possibly increased consumer education is a preferred alternative.

The results presented in this paper should be interpreted with caution. This study was intended as a preliminary examination of an area of study which was felt to have promise as a fertile area for future research. Certainly, not all variables which might be hypothesized to influence per capita rates of beverage alcohol consumption have been included in this study. Also, those variables which have been included, both dependent and independent, may suffer from problems of definition.

Despite the shortcomings of this preliminary study, it is possible to draw some interesting observations from it. It is possible also to regard this study as a 'hypothesis generating' exercise, important in the questions which it raises, rather than in those which it answers.

From the point of view of the authors, this study was designed to explore primarily the question of whether increased governmental regulation of advertising might be expected to contribute to a reduction in per capita consumption of beverage alcohol products. Certainly, the preliminary results presented herein do not permit a conclusive answer to this question, although they would suggest that regulation of advertising does not have a significant effect on per capita consumption, at least of beer and wine.

Additional research in this area is planned to explore more deeply the question of whether increased government regulation is necessarily the most valid approach to the solution of certain perceived social problems.

## References

- Robert Freed Bales, "Cultural Differences in Rates of Alcoholism", Quarterly Journal of Studies on Alcohol, 6, 1946, 480-499.
- Don Cahalan, Ira H. Cisin and Helen M. Crossley, American Drinking Practices: A National Study of Drinking Behavior and Attitudes (New Brunswick, N.J.: Rutgers Center of Alcohol Studies, 1969).
- Canada, Final Report of the Commission of Inquiry Into the Non-Medical Use of Drugs (Ottawa: Information Canada, 1973).
- Ira H. Cisin, and Don Cahalan, "Some Correlates of American Drinking Practices", in Nancy K. Mello and Jack H. Mendelson (eds.), Recent Advances in Studies of Alcoholism: An Interdisciplinary Symposium (Rockville, Md.: National Institute of Mental Health, 1971).
- Gerald Globetti, "Approaches to the Control of Alcoholic Beverages in the United States", Journal of Drug Issues, 3(No. 3, 1973), 260-266.
- Marvin E. Goldberg, and Gerald J. Gorn, "The Influence of In-Store Marketing on the Sale of Liquor", Working Paper, Faculty of Management, McGill University, 1975.
- Health and Welfare Canada, "Alcohol Problems in Canada: A Summary of Current Knowledge", No. ERD-75-22, Ottawa: 1975.
- James Hulbert, "Applying Buyer Behavior Analysis to Social Problems: The Case of Drug Use", in Ronald C. Curhan (ed.), 1974 Combined Proceedings (Chicago: American Marketing Association, 1974).
- J. A. Johnson, "Canadian Policies in Regard to the Taxation of Alcoholic Beverages", Working Paper 73-24, Department of Economics, McMaster University, 1973.
- James A. Johnson, and Ernest H. Oksanen, "Socio-Economic Determinants of the Consumption of Alcoholic Beverages", Working Paper 74-02, Department of Economics, McMaster University, 1974.
- Marc Lalonde, "On Marketing the Malt", Addictions, 21 (Winter, 1974), 32-40.
- Marc Lalonde, A New Perspective on the Health of Canadians: A Working Document (Ottawa: Information Canada, 1975).
- Gary H. Miller, and Neil Agnew, "The Ledermann Model of Alcohol Consumption", Quarterly Journal of Studies in Alcohol, 35, 1974, 877-898.
- Ontario Youth Secretariat, "Youth and Alcohol", Toronto: 1976.
- Robert E. Popham, Wolfgang Schmidt and Jan de Lint, "The Prevention of Alcoholism: Epidemiological Studies of the Effects of Government Control Measures", Toronto: Addiction Research Foundation, 1971.
- Richard Schmalensee, The Economics of Advertising (Amsterdam: North-Holland Publishing Company, 1972).
- Wolfgang Schmidt, "A Note of the Effect of Lowering the Legal Drinking Age on the Consumption of Alcoholic Beverages", Toronto: Addiction Research Foundation, 1972.
- Wolfgang Schmidt, and Jan de Lint, "Alcohol Use and Alcoholism", Addictions, 18(No. 2, 1971), 1-15.
- John R. Seeley, "Death by Liver Cirrhosis and the Price of Beverage Alcohol", The Canadian Medical Association Journal, 83(December 24, 1960), 1361-1366.
- Julian L. Simon, "The Price Elasticity of Liquor in the U.S. and a Simple Method of Determination", Econometrica, 34(January, 1966a), 193-205.
- Julian L. Simon, "The Economic Effects of State Monopoly of Packaged-Liquor Retailing", Journal of Political

Economy, 74(April, 1966b), 188-194.

Reginald G. Smart, "The Effect of Self-Service Stores on the Purchase of Alcoholic Beverages", Toronto: Addiction Research Foundation, 1974.

Henry Wechsler, Harold W. Demone, Jr., Denise Thum and Elizabeth H. Kasey, "Religious-Ethnic Differences in Alcohol Consumption", Journal of Health and Social Behavior, 11(March, 1970), 21-29.

## ASSESSING THE POTENTIAL EFFECTS OF DIFFERENTIAL PRICE INCREASES ON GASOLINE USAGE

Richard C. Reizenstein, The University of Tennessee  
David J. Barnaby, The University of Tennessee

### Abstract

This study investigates consumer reaction to various levels of gasoline price increase in terms of purchase intentions. Consumer groups formulated on the basis of propensity to consume less gasoline are comparatively analyzed. Energy policies are suggested for reducing gasoline usage and are discussed in light of varying implications.

### Introduction

For most of the twentieth century, American consumers have been accustomed to abundant supplies of relatively inexpensive energy. Products and services using these seemingly unlimited sources of power have been marketed by U. S. industry, fostering an energy affluent life style.<sup>1</sup> This almost unconscious dependency on an unending energy supply was abruptly brought into focus by the oil embargo in late 1973. Suddenly, the prospect of energy allocation, rapidly rising prices of energy intensive commodities and shortages of products and services became a reality, fostering a situation which has the potential to radically alter the American way of life. The most dramatic effects of the oil embargo were short-lived (i.e., lines in gasoline stations, curtailment of Sunday gasoline sales and quantity limitations at the pump); however, the threat of similar and even more extensive measures loom ever present on the horizon.

### Relevant Energy/Marketing Literature

The importance of petroleum to the U. S. economy and its profound effects on our lifestyle make it imperative that consumer attitudes, behavior and characteristics in regard to this basic resource be studied and analyzed. It is the purpose of this research to provide insights into this important area by identifying segments with varying propensities to conserve gasoline and to describe various characteristics of these groups. This approach will hopefully enable such segments to be more easily accessed, so that conservers may be reinforced; in contrast, differential marketing tactics (pricing, promotion, etc.) may be formulated to promote increased conservation in heavy gasoline consumption groups.

Attitudinal and behavioral information related to energy has been disseminated by both public opinion organizations and academicians. Opinion Research Corporation and similar firms have been conducting periodic nationwide research since the OPEC embargo, initiated on behalf of the Federal Energy Administration and private industry (including petroleum companies) (Opinion Research Corporation, 1974).

Several studies have been published by academicians. Talarzyk and Omura primarily focused on describing consumer attitudes toward leisure-time activities in the context of the energy crisis. Findings indicated that

<sup>1</sup>As an example of the relative magnitude of U. S. Energy consumption compared to other Western industrialized nations, see Helmut Becker and David J. Fritzsche, "Energy Consumption and Marketing: A Comparison of German and American Life Styles," *Marketing: 1776-1976 and Beyond*, Proceedings of the 1976 American Marketing Association, Kenneth L. Bernhardt, Editor, Series #39, pp. 527-532.

those respondents satisfied with their leisure-time activities felt that, while their lifestyles would be affected by the energy shortage, their non-energy intensive leisure activities would continue to provide increased satisfaction. (Omura and Talarzyk, 1974; Talarzyk and Omura, 1975). Potential, rather than existing, consumer attitudes were the object of a later study by Hass, Bagley and Rogers. Results of the study indicated that increments in the perceived noxiousness or severity of an energy crisis increased intentions to reduce consumption (Hass, Bagley, and Rogers, 1975). In contrast to these cross-sectional analyses, Barnaby and Reizenstein conducted a longitudinal study of consumer attitudes toward and perceptions of the energy situation. Measures of heat, air conditioning and gasoline usage were incorporated, as was an assessment of respondent attitudes toward air pollution and pollution/energy trade-off considerations (Barnaby and Reizenstein, 1975; Reizenstein and Barnaby, 1976).

As can be seen above, some attention has been given to assessing consumer attitudes toward the energy shortage as well as gasoline usage behavior. A further key variable which has been the focus of considerable speculation is consumer reaction to potential dramatic increases in price at the pump. During the years of the Ford administration, various pricing approaches to the control of gasoline consumption were debated. These focused primarily on the deregulation of the price of gasoline but have never been implemented, given projected adverse public response and the shift in national policy brought about by the election of President Carter.

Opponents to deregulation, such as Senators Abourezk, Jackson and Metzenbaum, have hypothesized that deregulation would result in an eight cent per gallon increase in price at the pump. These individuals maintain that such an increase in a necessity like gasoline would have maximum impact on those who could least afford it: lower and middle income customers who must either use public transportation, drive long distances to work or shop, or use gasoline as an integral requirement of their occupation (i.e., traveling salesmen, independent truck operators, etc.) (Bryon, 1977).

In contrast to the above viewpoint, there is evidence to indicate that decontrol of gasoline prices would not lead to significantly higher prices at the pump in the long run; in fact, Parisi asserts that deregulation would have little long run pricing effects (Parisi, 1977). There is, further, little data to indicate that the price of gasoline is sufficiently elastic that an abrupt increase would lead to a substantial diminution in demand. In fact, there has been conjecture that gasoline may be price inelastic within certain limited ranges. Such an argument can be approached from two different perspectives: 1) Bruce-Briggs contends that even given sharp increases in gasoline price, fuel still remains relatively cheaper than in the early 1960's as a percent of U. S. family income and that 1955 price levels relative to income would be maintained for a family of median income level with gasoline selling at \$.95/ gallon (Bruce-Briggs, 1974). Hence, substantial gasoline conservation would not take place without sudden large increases in price (perhaps to a dollar per gallon or more) and 2) Empirical data collected by Willenborg and Pitts indicated consumer intentions to modify gasoline consumption were high during periods of rapid gasoline price escalation. With the leveling off of the upsurge, however, the study showed

that consumers appeared to rapidly adapt to higher fuel prices and revised their purchase intentions upward to former consumption levels. The authors, therefore, concluded that a relative short-term inelasticity of overall demand for gasoline existed and suggested (similar to Bruce-Briggs) that only a dramatic increase in price of perhaps 100% would be an effective conservation mechanism (Willenborg and Pitts, 1977).<sup>2</sup>

In light of the uncertainties with respect to market reaction to gasoline price increases, this research attempts to identify and describe segments with varying conservation propensities at different price levels. This approach is viewed as a possible first step toward assessing the impact of incremental advances in the price of gasoline on consumer usage. It is hoped that this type of research might yield data which could provide insights into the advisability of implementing a pricing policy as a means of conserving scarce gasoline resources within unique consumer segments.

#### Methodology

After the OPEC oil embargo had been lifted and sufficient time had elapsed to enable consumers to adapt to a less crisis-oriented view of energy consumption, a mail survey was sent to 922 residents of three medium-sized Southeastern cities, selected by means of a systematic random sample using city directories. This questionnaire had as its primary thrust the measurement of consumer attitudes and behavior with regard to energy usage, with attitudes toward environmental issues, especially air pollution, being of secondary interest. Usable responses were obtained from 390 subjects in the sample, a response rate of nearly 42%. Demographic comparisons with updated census statistics showed the sample to be generally representative of the population with the exception of a slightly upscale skew in the education and income distributions.

#### Analysis

A variety of attitudinal and behavioral data was collected in the survey. This included energy- and pollution-related attitude and activity measures, information source variables and measures of consumer intention with regard to gasoline usage based on hypothetical per gallon increases in the price of gasoline. These variables were used to focus on identifying consumer groups exhibiting different gasoline conservation intentions with respect to the varying per gallon price increments. Once groups were formulated, multiple discriminant analyses were employed to describe segments in terms of their energy-related attitudes, use of specific information sources and selected demographic characteristics.

Prior to exploring the extent of the relationship between the set of predictors listed above and the price sensitivity groups, however, a factor analysis of the large set of attitude variables was performed to yield a reduced set of factors. These composite variables were included as independent variables in multiple discriminant analyses on groups structured according to the number of gallons per week of projected reduction in gasoline consumption.

<sup>2</sup>It is interesting to note that while the price of gasoline continues to be driven upward by the OPEC cartel, it is in actuality rising no more rapidly than the rate of U. S. inflation. In fact, it can be seen that the recent price of gasoline at the pump as measured in 1967 constant dollars is slightly lower now than it was in 1974 during the embargo.

#### Results

Respondents to the questionnaire (n = 390) were asked to indicate how much they would be willing to reduce current gasoline consumption for each of nine possible five-cent increments in price per gallon at the pump. The data generated by this question are presented graphically in Figure A, with hypothesized price increases displayed along the horizontal axis and respondents grouped cumulatively along the vertical axis according to that percentage projecting conservation of a defined number of gallons of gasoline per week.

These results indicate that increases of up to \$.15 per gallon at the pump resulted in 35% of the sample indicating that they would reduce gasoline consumption by at least one gallon of gasoline per week (9 percent indicated they would reduce consumption by at least 7 gallons per week). Further increases in gasoline price up to \$.45 per gallon produced further decreases in consumption by up to 48% of the respondents (18 percent would reduce 1-6 gallons per week and 30 percent would reduce 7-12 gallons per week); however, 52 percent of the survey participants remained price inelastic, being unwilling to reduce gasoline usage under any of the price increases offered.

Although some caution must be exercised in accepting these data at face value given the propensity of consumers to sometimes behave in a manner at variance with their stated intentions, historical data seem to support the potential validity of the intentions measures cited in the survey.<sup>3</sup> Examination of such data indicates that the only real reduction in U. S. demand for petroleum occurred in early 1974, shortly after the only abrupt increase in the real dollar price of gasoline at the pump. Thus, historical backcasting seems to verify that at least a significant portion of the population is price elastic given a dramatic real increase in the cost of gasoline. (Note that in Figure A, a projected \$.10 per gallon increase causes only 8% of the sample to state that they are willing to reduce gasoline consumption by 7-12 gallons per week, while a \$.45 per gallon increase similarly affects 33% of the sample.)

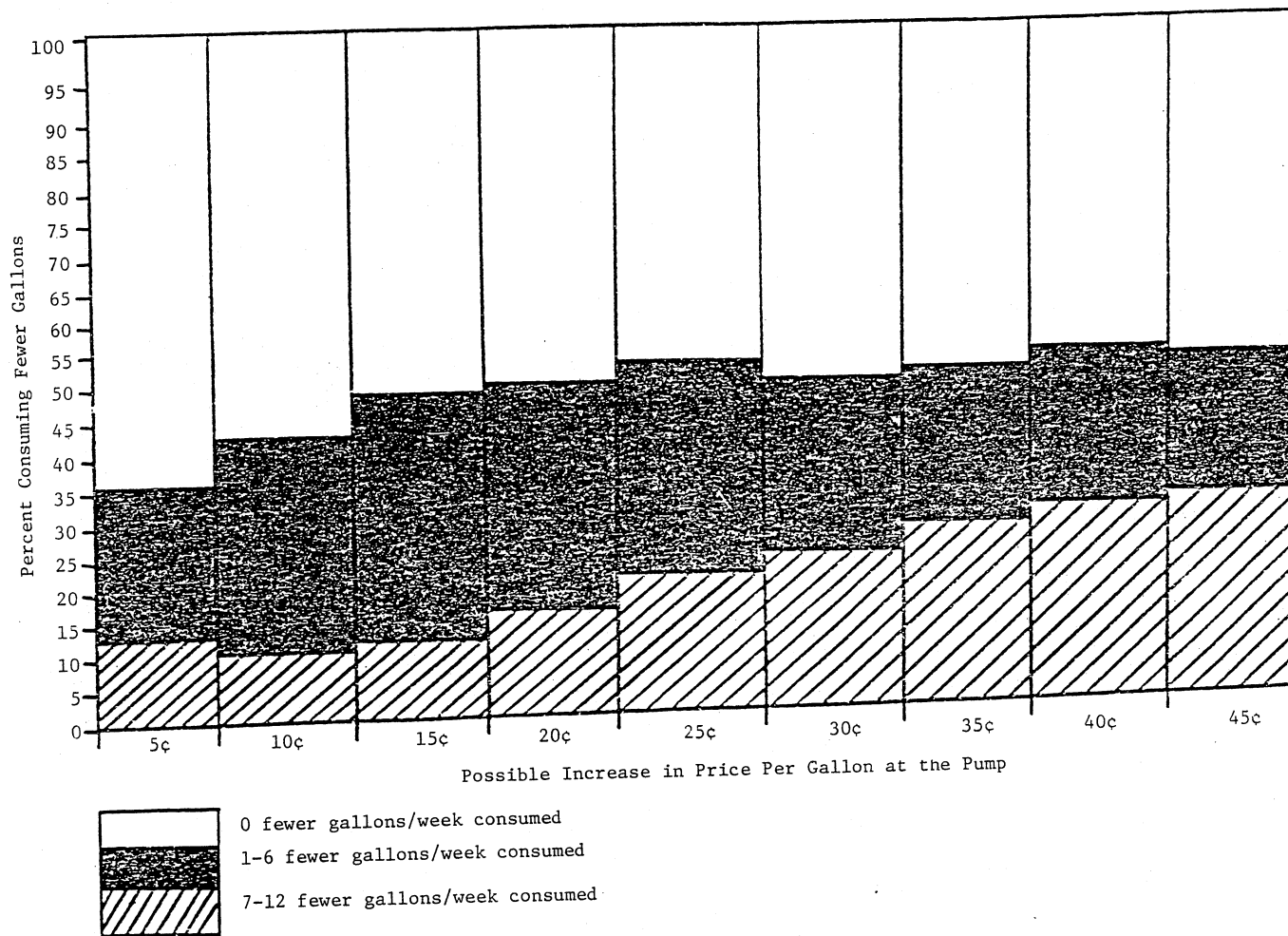
It thus seems evident from this analysis, at least according to intentions measures, that there is a sizable group of respondents which appears price insensitive (price inelastic), while others demonstrate varying degrees of price elasticity. These results agree in part with the findings of Willenborg and Pitts showing inelasticity of demand within limited increments in price. In addition, however, the results from Figure A provide some evidence of price elasticity within smaller respondent groupings. This finding suggests that the entire gasoline market response should not be characterized as elastic or inelastic, but rather may include consumer segments possessing differing gasoline demand functions (and, conversely, different conservation functions). In order for operational promotional strategies to be derived from this conclusion, further indepth analysis of these segments is mandatory. Group differences must be analyzed to determine the existence of segment characteristics (in this case attitudes, demographics and information sources) that are sufficiently unique to warrant the formulation of market-oriented conservation strategies and tactics for specific groups.

The objective of attempting to identify distinct gasoline conservation segments based on various consumer characteristics was met by using multiple discriminant analysis (MDA) to maximally separate groups according to a set of

<sup>3</sup>Since the scale employed has many of the properties of a purchase probability scale, this augments the probability of predictive accuracy due to the greater sensitivity of the PPS approach in distributing responses along the intentions continuum (Juster, 1966).

FIGURE A

THE EFFECTS OF POSSIBLE GASOLINE PRICE LEVELS ON CONSUMPTION



predictor variables. First, a factor analysis of the 45 AIO variables yielded a set of 15 factors that served as one portion of the predictors used to attempt to describe the various gasoline conservation groups.<sup>4</sup> In addition, both media and personal information source utilization variables and demographic measures were included as potential group descriptors.

Table 1 describes the results of SPSS multiple discriminant analysis of three consumer segments espousing different gasoline conservation intentions: reduce consumption by 0, 1-6 and 7-12 gallons per week, assuming a \$.10 per gallon increase. In total, 46 different predictors, including AIO factor scores, information sources and demographics, were incorporated in the analysis. The multivariate F-ratio, significant at the .005 level, provides support for overall group discrimination; however, inspection of significant univariate F-ratios of the 46 predictors shows only 6 significant variables.<sup>5</sup>

Examination of the significant variable means for each of the three groups shows relatively little separation. Caution must therefore be exercised in interpreting group differences since only relative comparisons are in order

<sup>4</sup> For a more detailed exposition of the factors generated and those AIO's loading on specific factors see David J. Barnaby and Richard C. Reizenstein, "Analyzing Gasoline Consumption After the Oil Embargo: A Market Segmentation Approach," a paper presented at the ORSA/TIMS Conference, Miami, Florida, November 3-5, 1976.

(see particularly factor score mean values); however, the data do provide insights into the attitudinal and demographic composition of the three groups as well as information source importance. These data, presented in Table 1, are comparatively described for the three gasoline consumption segments in Table 2.

An immediately noticeable characteristic of Group 1 is its large size (n = 240) as compared to either of the other two groups. The significant attitudinal variables seem to indicate a lack of concern with energy- or pollution-oriented issues with respect to the other groups. In contrast, Group 3 appears to have more of an activist orientation in regard to energy/environmental issues, while Group 2 may be viewed as energy aware. This latter group has realized the potential impact of the energy shortage (although not to the same degree as Group 3); however, this segment perceives energy problems as more immediately important than pollution abatement.

<sup>5</sup> It is important to note that the greatly unequal group sizes (immediately evident in Table 1 across price increase levels) are a major consideration in this analysis, given the disproportionate group classification likelihoods associated with the unequal categories. The disproportionately small size of Group 3 (n = 25) precluded a split-half analysis to verify the reliability of the six significant predictors. Reclassification via the discriminant functions, however, always exceeded chance probability of group membership.

TABLE 1

DISCRIMINANT ANALYSIS OF THREE GASOLINE USAGE REDUCTION SEGMENTS GIVEN A 30 CENT/GALLON FUEL PRICE INCREASE<sup>a</sup>

Variable	F-Ratio <sup>b</sup>	Group 1	Group 2	Group 3	
		Reduce Consumption by 0 gallons (n = 240)	Reduce Consumption by 1-6 gallons (n = 125)	Reduce Consumption by 7-12 gallons (n = 25)	
No.	Name	$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_3$	
20	Information received by mail	6.20	.66	.86	1.28
31	Number of automobiles owned	3.10	1.82	1.60	1.72
09	The energy shortage will hurt our economy	3.74	3.14	3.32	3.60
11	Fuel alternatives to gasoline will power future automobiles	3.98	2.22	2.38	1.87
03	Energy problems are more important than pollution abatement	3.14	1.20	1.27	.67
18	Radio	4.88	1.53	1.63	2.24

<sup>a</sup>SPSS Multiple Discriminant Analysis of three groups using 46 predictor variables and a sample of 390 produced results shown in the table. Multivariate F-Ratio significant at the .005 level.

<sup>b</sup>Univariate F-Ratios with 2 and 387 degrees of freedom; F.05, 2, 120=3.07.

Media information source comparisons indicated that Group 1, the least concerned segment, also relies least on the two significant media information sources (direct mail and radio). In addition, this group includes the greatest average number of two car families. Thus, it

would appear that the least concerned segment may also be the least information-oriented segment and the least willing to give up fuel for their two cars. In comparison, although Group 3 still has a preponderance of two car families, it appears much more willing to reduce fuel consumption significantly. The fact that this group is the most oriented toward information sources, such as direct mail and radio, may be an important contributing factor toward this orientation. Given the size and orientation of Group 1 as contrasted to Group 3, it would seem as though a \$.10 per gallon increase in the price of gasoline may not be an incentive to change gasoline consumption patterns in a market segment that has the least concern for the energy situation. It would appear that only those with a greater energy awareness, who are more information oriented, are sufficiently sensitive toward the problem to be willing to sacrifice gasoline consumption.

Similar analyses were performed for \$.20, \$.30 and \$.40 per gallon price increases to determine attendant shifts in consumption and segment composition. Only the results of the \$.30 price increase analysis will be shown, however, as it is at this level that the most meaningful group changes are evidenced. Beyond this point, segment composition and descriptions do not seem to change appreciably.

The data from the discriminant analysis for the \$.30 per gallon fuel price information are presented in Table 3. These data are converted into verbal comparisons of the three gasoline conservation groups in Table 4. Immediately noticeable in all of these analyses is the change in size of the three groups as compared to the \$.10 price increase. Groups 1 and 2 have been reduced in total by 61 respondents; Group 3 has grown commensurately larger. (Additional price increases within the limits investigated in this study did not result in noteworthy changes

TABLE 2

## DESCRIPTION OF GASOLINE CONSERVATION GROUPS GIVEN A 10 CENT/GALLON FUEL PRICE INCREASE

Group 1	Group 2	Group 3
Reduce Consumption by 0 Gallons/Week (n = 240)	Reduce Consumption by 1-6 Gallons/Week (n = 125)	Reduce Consumption by 7-12 Gallons/Week (n = 25)
<u>Consumer Attitudes</u>		
Agree least that the energy shortage will hurt our economy.	Attitude toward the impact of the energy shortage on our economy falls between that of Group 1 and Group 3.	Agree most that the energy shortage will hurt our economy.
Attitude toward alternatives to gasoline as future automotive fuel falls between that of Group 2 and Group 3.	Agree most that fuel alternatives to gasoline will power future automobiles.	Agree least that fuel alternatives to gasoline will power future automobiles.
Attitude toward relative importance of energy problems and pollution abatement is between that of Group 2 and Group 3.	Agree most that energy problems are more important than pollution abatement.	Agree least that energy problems are more important than pollution abatement.
<u>Media Information Sources</u>		
Information received by mail no help as information source.	Information received by mail little to no help as information source more than with Group 1	Information received by mail little to some help as information source more than with Groups 1 or 2.
<u>Demographic Characteristics</u>		
On the average, more two car families than single car families in this group	On the average approximately an even mix of two car and single car families in this group.	On the average, more two car families than single car families in this group, but fewer than in Group 1.



in group size.) These results reinforce the previous contention derived from Figure A and supported by the data in Tables 1 and 2 that a sizable number, 52% of the respondents in Group 1, continue to be price inelastic with regard to gasoline consumption. Given the above results with the attendant shift of some respondents to more conservation-oriented segments, a reexamination of the composition of the three gasoline conservation groups at the \$.30 level is warranted.

A comparison of group sizes shown in Table 3 with the parallel analyses reported in Table 1 demonstrates the shift toward greater gasoline conservation with the increase in the membership of Groups 1 and 2. The increase in price seems to have refocused the outlook of those who do not intend to reduce their gasoline consumption (Group 1) at any of the proposed levels of price increase. Whereas this segment formerly felt that the energy shortage would not appreciably hurt the economy, the potential of a \$.30 gallon fuel price increase has resulted in a greater degree of pessimism and skepticism with respect to the effects of the energy situation on everyday life. For example, those in this segment now believe more than other respondents that the energy crisis will create personal hardships, while maintaining the rather skeptical view that the energy shortfall was precipitated by the oil companies. In addition, they are least confident in the fuel efficiency of the new automobiles.

TABLE 3

DISCRIMINANT ANALYSIS OF THREE GASOLINE USAGE REDUCTION SEGMENTS GIVEN A 30 CENT/GALLON FUEL PRICE INCREASE<sup>a</sup>

No.	Variable Name	F-Ratio <sup>b</sup>	Group 1	Group 2	Group 3
			Reduce Consumption by 0 Gallons (n = 203) $\bar{X}_1$	Reduce Consumption by 1-6 Gallons (n = 101) $\bar{X}_2$	Reduce Consumption by 7-12 Gallons (n = 86) $\bar{X}_3$
43	Age	7.26	4.77	4.41	4.11
44	Education	8.09	3.61	4.28	4.01
10	The energy crisis will create personal hardships for our citizens	4.30	2.61	2.40	2.23
13	The newer automobile engines are highly fuel efficient	5.03	1.53	1.67	1.77
02	The energy crisis was created by the oil companies	3.55	1.62	1.29	1.33
35	No. of children 0-5 years	3.33	.22	.24	.42
42	Mobility	3.18	1.61	1.76	1.94

<sup>a</sup>SPSS Multiple Discriminant Analysis of three groups using 46 predictor variables and a sample of 390 produced results shown in the table. Multivariate F-Ratio significant at the .005 level.

<sup>b</sup>Univariate F-Ratios with 2 and 387 degrees of freedom; F .05, 2, 120 = 3.07.

It is also of interest to note that the discriminant analysis at the \$.30 level yields no information source variables as significant predictors, but does provide a more distinct demographic profile of the three groups. As the hypothetical levels of gasoline price increases rise, more socioeconomic rather than information source and attitudinal variables emerge as significant discriminators. Consumer attitudes appear to have a lesser impact at the higher gasoline price levels when contrasted to demographic and situational constraints which impose boundaries on the individual's ability to react to a changing energy environment. This phenomenon can be noted by examining the Group 1 demographic descriptors in Tables 3 and 4. In contrast to the \$.10 price increase

level, where car ownership surfaced as the only significant demographic characteristics, the \$.30 price increase profile shows the average Group 1 respondent to be older (and thus with fewer young children), less educated and less mobile than his Group 2 and Group 3 counterparts. These variables may contribute to the attitudinal rigidity described previously.

In contrast, the most conservatively oriented segment, Group 3, has enlarged considerably when compared to the \$.10 per gallon price increase group (22% as compared to 6% of the total response). The attitudes differentiating this from the other groups seem to have been somewhat re-directed by the additional respondents who have shifted to Group 3. This segment formerly agreed most that the energy crisis would hurt the economy; now, under the higher potential price increase, it appears to be the least apprehensive of the three groups, agreeing least that the energy crisis will cause personal hardships. Perhaps this attitude can be attributed in part to the greater degree of confidence in the fuel efficiency of the newer automobile engines. It is possible that this segment feels that the newer model autos will enable them to save a sufficient amount of fuel per mile driven, such that they can continue to drive nearly as much as they had driven prior to the oil embargo and still save an appreciable amount of gasoline. This interpretation would also partially explain the more optimistic attitude of this group in relation to the energy crisis not causing personal hardship. (This conclusion is consistent with the results of the study by Willenborg and Pitts.)

Socioeconomic variables continue to play an important role as Group 3 descriptors. In contrast to the older, less mobile Group 1, Group 3 members are, on the average, less than 40 years of age with almost half having young children. They are, in addition, somewhat more educated and much more mobile than Group 1.

Predictably, the mean descriptor values of Group 2 seem to fall between Groups 1 and 3 on all dimensions judged significant in the discriminant analysis with the exception of one variable: education. This segment is marginally the most educated of three groups, a factor which may have contributed to some conservation intentions but which by itself is not sufficient to have stimulated maximum reduction in gasoline usage.

Conclusions

A major use of market segmentation has traditionally been to analyze markets with respect to the consumption of goods and services. This study has demonstrated the versatility of a segmentation approach by applying it to potential conservation behavior. In addition, multivariate analyses of the different data sets have provided insights into the profiles of conservation group members at different gasoline price increment levels. These profiles were even further amplified by the use of several different types of descriptors: attitudinal, personal and media information source, and socioeconomic. Thus, the combination of the segmentation concept, the multivariate approach to profiling, and the incremental depth inherent in the descriptor variables, has provided insights into both segment profiles and marketing approaches designed to match the potential conservation orientations of the relevant groups.

In general, the analyses supported a division of the respondents into two broad gasoline conservation segments: 1) those who were basically price inelastic, who were non-conservation oriented and who did not intend to voluntarily change gasoline conservation practices either as a result of promotional appeals and/or pricing mechanisms; and 2) those who possessed varying degrees of price inelasticity and a conservation orientation, who might well be responsive to some combination of promotional appeals

TABLE 4

## DESCRIPTION OF GASOLINE CONSERVATION GROUPS GIVEN A 30 CENT/GALLON FUEL PRICE INCREASE

Group 1 Reduce Consumption by 0 Gallons/Week (n = 203)	Group 2 Reduce Consumption by 1-6 Gallons/Week (n = 101)	Group 3 Reduce Consumption by 7-12 Gallons/Week (n = 86)
Agree most that the energy crisis will create personal hardships for our citizens.	Attitude toward the energy crisis as a cause of personal hardships for our citizens falls between that of Group 1 and Group 3.	Agree least that the energy crisis will cause personal hardships for our citizens.
Agree least that the newer automobile engines are highly fuel efficient.	Attitude toward the greater fuel efficiency of the newer automobile engines falls between that of Group 1 and Group 3.	Agree most the the newer automobile engines are highly fuel efficient.
Agree most that the energy crisis was created by the oil companies.	Agree least that the energy crisis was created by the oil companies.	Attitude toward the role of the oil companies in creating the energy crisis falls between that of Group 1 and Group 2.
Average age in mid-40's.	Average age in early 40's.	Average age less than 40.
High school graduates, some with 1-3 years college.	1-3 years college, some with four year college degrees.	1-3 years college.
Few have a young child. (0-5 years old)	Few have a young child. (0-5 years old)	Nearly half have a young child. (0-5 years old)
More than half have moved more than once during the last five years, on the average.	The great majority have moved more than once during the last five years, on the average.	Nearly all have moved an average of twice during the last five years.

and/or pricing mechanisms. Moreover, within this latter group, a further breakdown identified subsets of consumers in regard to their conservation propensities and their gasoline consumption practices. At different price increase levels (\$.10 vs. \$.30), a definite shift occurred from discrimination among groups based on information source and attitudinal variables to discrimination based on attitudinal, situational and demographic variables. This seems to indicate that with greater price increases, crystallization of groups will occur and situational variables will become more critical in shaping conservation intentions. (This contention is supported by the absence of information source variables and the emergence of more situational and demographic discriminators in the \$.30 analysis.) This also may imply that promotional appeals, even when targeted toward more conservation oriented segments, may be less effective in changing attitudes with greater levels of price increase, given socioeconomic constraints.

Although the results indicate that conservation intentions can be elicited from respondents similar to those in Groups 2 and 3, it is evident that the largest segment remains price inelastic regardless of the price increase within the \$.05-\$.45/gallon price range under investigation. Thus, promotional and pricing mechanisms dependent on voluntary consumer response appears an ineffective alternative for this group. It seems as if some form of mandatory compliance legislated at the national level may be necessary to effect substantial conservation across all segments. Over and above such uniform requirements of all citizens (such as the Carter Administration's proposed gas guzzler tax, and the required 1985 gasoline mileage standards for automobile manufacturers), market related appeals could potentially impact more conservation-oriented segments (such as Groups 2 and 3). Thus, necessary conservation could be accomplished for the most part via legislation, relying on voluntary measures for less critical contributions to gasoline conservation. Finally, legislated appeals would not only be more uniform

in their overall requirements of the population, but would also undoubtedly be more successful in accomplishing necessary conservation objectives than would a strictly voluntary program to which only a portion of the population would respond.

## References

- Barnaby, David J. and Richard C. Reizenstein, "Attitudes Toward Consumption: Segmenting the Gasoline Market," *Advances in Consumer Research*, Volume 3, B. B. Anderson, Editor, (Cincinnati: Association for Consumer Research, 1975).
- Bruce-Briggs, B., "Gasoline Prices and the Suburban Way of Life," *The Public Interest*, No. 37, Fall, 1974, pp. 131-136.
- Byron, Christopher, "Yes, There Is An Energy Crisis," *Time*, Vol. 110, No. 15, October 10, 1977, pp. 61-62.
- Hass, Jane W., Gerrold S. Bagley, and Ronald W. Rogers, "Coping with the Energy Crisis: Effects of Fear Appeals Upon Attitudes Toward Energy Consumption," *Journal of Applied Psychology*, 1975, Vol. 60, No. 6, pp. 754-756.
- Juster, Thomas F., *Consumer Buying Intentions and Purchasing Probability*, (New York: National Bureau of Economic Research, 1966).
- Omura, Glenn S. and W. Wayne Talarzyk, "Relationships Between Consumers Shopping and Leisure Time Activities and Their Attitudes Toward the Energy Crisis: A Cross Sectional Study," in *Advances in Consumer Research*, Vol. 2, Mary Schlinger, Editor, (Chicago: Association for Consumer Research, 1974), pp. 803-816.
- Parisi, Anthony, "The Myth of Gasoline Price Controls," *Business Week*, No. 2466, January 17, 1977, p. 25.

Reizenstein, Richard C. and David J. Barnaby, "An Analysis of Selected Consumer Energy Environment Trade-Off Segments," in Marketing: 1776-1976 and Beyond, Kenneth L. Bernhardt, Editor, (Chicago: American Marketing Association, 1976).

Reizenstein, Richard C. and David J. Barnaby, "The Impact of the Energy Crisis on Consumer Willingness to Spend to Reduce Air Pollution," in Ecological Marketing, Karl Henion II and Thomas C. Kinnear, Editors, (Chicago: American Marketing Association, 1976).

Talarzyk, Wayne W. and Glen S. Omura, "Consumer Attitudes Toward and Perceptions of the Energy Crisis," in Marketing Contributions to the Firm and to Society, Ronald C. Curhan, Editor, (Chicago: American Marketing Association, 1975), pp. 316-322.

Willenborg, John F. and Robert E. Pitts, "Gasoline Prices: Their Effect on Consumer Behavior and Attitudes," Journal of Marketing, Vol. 41, No. 1, January, 1977, pp. 24-31.

F. Robert Dwyer, University of Cincinnati

## Abstract

A brief review of the drug compliance literature is presented in order to frame this and perhaps other primary research efforts. A clinical study is outlined and two linear models are used to test the previously examined factors of compliance in conjunction with an untried variable, health locus of control. Third, the jackknife validation procedure is described and demonstrated to highlight its particular relevance in compliance and other sample-constrained research settings.

## Drug Defaulting

There have been nearly 70 studies on patients who fail (intentionally or accidentally) to take their medications "as directed." In these studies, a variety of methods have been used to detect the drug defaulter: (1) interrogation, (2) pill counts, (3) stool and urine markers, (4) drug detection in blood or urine, and (5) failure to dispense or refill the prescription (Blackwell, 1972). Doctor intuition in detecting drug defaulters has been very unreliable; error rates are slightly better than chance, and is therefore seldom used in studies of compliance (Caron and Roth, 1968).

Boyd et al. (1974) have reviewed 35 patient compliance studies and have found defaulting rates from 4% to 90%, depending on the drug, patient group and method of assessment. On the latter point these authors computed weighted averages of noncompliance for the three primary methods of detection: urine tests, dosage counts, and patient self reports. While the mean rates of defaulting were comparable: 36%, 47%, and 38%, respectively, for each type of measure there is no indication of the convergent validity of the measures.

The consequences of overdoses, omissions, and improper mixing of prescription drugs involve varying health hazards for the defaulter. Patients who suffer from arthritis and default from anti-inflammatory drugs are at worst trading-off their own pain against the inconveniences of continuous medication. In a curative situation, a study by Leisty and MacAuley (1966) showed over 75% of the patients defaulting from the prescribed number of doses of oral penicillin for their streptococcal pharyngitis had positive throat cultures. All compliant patients showed negative cultures.

If a physician is unaware that a hypertensive patient is not taking the medication according to directions and sees that the blood pressure is still elevated, the physician may prescribe larger doses of the same agent(s) or may prescribe a more potent anti-hypertensive medication (Hussar, 1975). On a similar note, patient non-compliance may play a part in distorting the controlled evaluation of drug therapy.

In addition to depriving the patient of the anticipated therapeutic benefits and possibly resulting in a progressive worsening of the condition being treated, the consequences of drug defaulting include the economic wastage involved and the hazard to health posed by cupboards stocked with unused or unidentified medications (Blackwell, 1972).

## Factors in Noncompliance

The extent, nature, and consequences of drug defaulting have highlighted its significance from both a patient and

societal perspective. Logically, numerous studies have attempted to uncover patient reasons and circumstantial factors associated with drug defaulting. While the findings have been somewhat inconclusive on the demographic and psycho-social correlates of compliance, conclusions on the impact of certain medication and treatment factors appear more consistent. Even on this latter point, caution is required since findings of no statistical significance on tested variables are often not reported and in no case to date have researchers validated their method or findings on a new set of data.

## Treatment Factors

After an extensive literature review, Blackwell (1973) concludes:

Patients with prolonged conditions are clearly prone to lapses in compliance, especially when the treatment is prophylactic or suppressive (as with malaria), when the condition is mild or asymptomatic (in anemia of pregnancy) or when the consequences of stopping therapy may be delayed (for instance, as in epilepsy or schizopenia). (p. 241)

The literature is replete with studies showing discontinuation of oral antibiotics because patients "feel better," yet the infection may not be under complete control.

Brook and associates (1971) showed that when noncompliance is likely to lead to an immediate and/or severe relapse, the patient is more likely to comply. The cardiac patients in their study were more adherent to their regimen of digoxin and diuretics than they were to their potassium supplements.

The effect of multiple medications on patient compliance has been widely studied. Clinite and Kabat (1969), Madden (1973), Kellaway and McCrae (1975), and Malahy (1966) found that compliance decreased among patients taking three or more medications. Noncompliance is also encouraged by multiple doses of the same drug. The magnitude of this effect was demonstrated by Gatley (1968) when the number of defaulters (in a prospective trial) doubled as the number of tablets was increased from one to four per day.

It seems logical that a patient who experiences side effects is less likely to comply than the patient with no side effects. This suspicion is confirmed by a number of studies. Wynn-Williams and Artis (1958) found that PAS (Para-amino-salicylic acid) has produced gastrointestinal upset in many more defaulters than in non-defaulters. Hussar (1975) reports that a number of observers have noted that the development of impotence is a major reason for male patients to discontinue taking anti-psychotic medications.

A treatment dimension related to complexity, yet requiring special attention, is life style change mandates that can accompany prescription medications. For example, it is desirable for patients being treated with tranquilizers or sedatives to avoid excessive consumption of alcoholic beverages because of the increased likelihood of an excessive depressant effect. Many patients choose to cope with this mandate by simply not taking their drug (Hussar, 1975).

## Patient Characteristics

Studies which have explored the psycho-social and demographic correlates of noncompliance have generally shown mixed results. However, no comprehensive studies of the entire patient population have ever been conducted. Therefore, research conclusions are constrained by the limited age range, economic status range, etc., of the accessible geographic, disease subpopulation studied.

Bakker and Dightman (1964) administered a battery of personality tests to women taking oral contraceptives. Those women who failed to take the Pill regularly were found to be more immature, irresponsible, and impulsive than their compliant counterparts. In addition, their personality profiles deviated more from their husband's than did those profiles of women who took the pill regularly. It is likely that some other personality dimensions would be associated with noncompliance for other medications.

Blackwell (1972) reports that some psychoanalytic writers have suggested that pills and capsules are symbolic representatives of the breast and penis. Little evidence in support of this notion is reported, however. More common in the psychiatrist's practice are patient paranoid delusions that cause the schizophrenic to equate drugs with poison. Also, Hussar (1975) reports an increasing number of patients concerned about becoming drug-dependent. To avoid such a possibility and to prove to themselves that they are not dependent, they may interrupt or by some other means modify their treatment regimen.

As might be expected, the patient-physician relationship can affect compliance. Francis *et al.* (1969) examined parent attitudes toward the physician in compliance of pediatric outpatients and found that compliance was reduced when the mother perceived the doctor as unfriendly and insensitive to the child's condition or illness.

A handful of studies have shown positive correlations between compliance and the patient's perceptions of susceptibility to disease, seriousness of the condition and efficacy of the treatment (Sharpe, 1977). In an attempt to explain preventive health behaviors, these factors, plus a patient's analysis of the "benefits and barriers" of treatment and a "cue to action," have been formulated by Rosenstock (1966) into what has been labeled the health belief model. Its performance has been mixed in the preventive setting (Becker *et al.*, 1972; Oliver and Berger, 1978) while in a curative setting only its individual components have shown occasional significance (Sacket and Haynes, 1976).

Moving away from psychological correlates, Boyd *et al.* (1974) after collecting data on 380 prescriptions from 134 outpatients, made several conclusions on the effects of demographic variables -- sex, age, race, education, occupation, and social class -- on compliance. No differences in error rates were noted between races, sexes, education levels and the Hollingshead Index-defined social classes. Differences were identified by age groups, however. The 45-64 age group had the least number of errors per prescription. The authors' explanations of the age effects were:

1. Patients in the 25-44 age group are preoccupied in the working world and are not acutely health conscious.
2. The 45-64 age group is becoming more aware of their health while concurrently decreasing their other activities.
3. The 65 and over group is experiencing more complex health problems combined with decreasing ability to care for themselves.

Malahy's (1966) analysis of demographic effects was narrower than Boyd's. While the studies are consistent in their rejection of education effects, Malahy went on to conclude no age effects on compliance. The apparent discrepancy seems to be due to the limited range of Malahy's sample -- most patients were between the ages of 35 and 65.

Clinite and Kabat's (1969) study of 1,060 patients, however, suggests the question of age effects cannot be put aside. They noted that patients between 71 and 80 years had by far the lowest rate of noncompliance. It is possible that even Boyd's sample (n = 134) was not sufficiently broad to negate the variance impact of previously mentioned treatment factors associated with noncompliance and likely to covary with age. Examples include increased seriousness of disease and consequences of defaulting, multiple medications and complex treatments, and advancing senility.

Madden (1973) joins Malahy and Boyd *et al.* in finding an absence of education effects on compliance with antibiotic regimens. However, other studies have reported higher education levels for defaulters among tuberculosis patients and lower education standards for defaulters among pregnant women taking iron (Berry *et al.*, 1963; Porter, 1969).

Income seems to be directly related to compliance, as several studies have shown that for some families the money is not always available to get the prescription filled or refilled (Gurwich and Emmanuel, 1974).

In their study of the compliance rates of mental outpatients, Lipman *et al.* (1965) found black patients less compliant than whites. A later study of psychiatric patients by Hare and Wilcox (1967) found no difference in compliance between blacks and whites. Adding to the confusion of the effects of subculture is a study by Kellaway and McCrae (1975) in New Zealand. They found Polynesian patients more likely to make simple errors -- unintentional mistakes -- than European patients. On the other hand, European patients showed a greater tendency for noncompliance involving a positive decision with at least some degree of volitional intent to alter or interfere with the prescribed regimen.

It is logical to expect patients in hospitals to show greater compliance rates than outpatients. This assumption was verified by Hare and Wilcox who checked urine markers of in-patients, day-patients, and outpatients. Employing moderate criteria of interpretation of urine analysis, noncompliance rates were 19%, 39% and 48% for inpatients, day, and outpatients, respectively.

The effect of supervision of drug therapy may be reflected in the reported effects of marital status on compliance. A study of 220 chronically ill patients over 60 conducted by Schwartz *et al.* (1962) found married persons more compliant than patients living alone. Porter (1969) studied a broader population of patients on long term treatments, but also concluded that socially isolated patients tended to neglect their drugs.

The effects of a patient's previous experience with the same or a similar disease is unclear. In Ball's (1974) review of several studies he concludes experience has a positive impact on compliance. However, Lipman, *et al.* found experience to have a negative impact on compliance with meprobamate. In proper perspective, it seems experience would interact with the nature of the treatment and previous outcome on compliance.

## Comprehension of Regimen

In addition to all these factors of compliance,

Blackwell (1973) concluded on the basis of his own empirical evidence:

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 clearcut, though, as Malahy (1966) and Sackett *et al.* (1975) failed to find compliance effects from increased comprehension of treatment and illness.

Comprehension effects on compliance are of particular interest to policy makers as written, patient-oriented information may soon be dispensed with all or most prescription drugs.

#### Research Problems

Two basic administrative problems have confounded researchers in the compliance area. First, capturing an adequate sample size is constrained by the realm of experimenter's control and subject arrival rates. Due to the convenience of using the single researcher-affiliated health care site and/or coordination problems arising from multifacility studies, the former approach is most widely implemented. The single site constraint has a striking effect on the patient population studied. It means the researchers must study compliance in a patient population undergoing a variety of treatment programs or must study compliance in patients with a prevalent condition. The former thrust increases both the total variance and the proportion of variance from disease and treatment factors. The latter usually means a lengthy data collection period and a focus limited to such prevalent conditions as strep throat and hypertension.

The above-mentioned administrative problems tend to magnify the seriousness of analysis difficulties encountered. Typically computerized data searching techniques, stepwise regression or discriminant analysis, are utilized. These procedures are useful for exploratory research. However, because of their tendency to capitalize on spurious correlations in the data, they often uncover "significant" relationships (that often have intuitive appeal) that are, in fact, random. Before such findings can be labeled conclusive they must be tested on a new set of data.

This call for validation merely echoes the cries many consumer researchers have previously sounded (Sheth, 1977; Etgar, 1977). The most frequently recommended and utilized approach has been the holdout method. Here the data are randomly split; one half is used for model estimation, the other tests its stability or predictive validity. This procedure is infeasible, however, when costs and administrative constraints preclude an adequate sample size. Statistics from each data set may be biased in proportion to  $1/n$  and therefore quite large when the sample size is small.

The following section describes a more efficient validation method, the jackknife, and demonstrates its viability in a study of compliance.

#### The Jackknife

Nearly 20 years ago Quenouille (1956) and Tukey (1958) proposed a methodology for reducing estimation bias in small samples. The technique, labeled the jackknife by Tukey for its versatility as a statistical tool, offers

ways to set sensible confidence limits in complex situations. An excellent detailed discussion of the technique is provided in Mosteller and Tukey (1977) and Miller (1974).

Briefly, the crux of the jackknife is to divide the data into  $g$  groups of size  $h$ ,  $n = gh$ , and assess the effect of each of the groups by examining the effect on the sample statistics that results from omitting that group.

This is accomplished by making  $Y(j)$  the results of the complex calculation on the portion of the sample that omits the  $j$ th subgroup; estimates are made on a pool of  $(g - 1)$  subgroups. By subtracting  $Y(j)$  times  $(g - 1)$  from  $g$  times  $Y_{all}$ , the corresponding result for the entire sample pseudo-values  $Y_j^*$  are obtained. Equivalently:

$$Y_j^* = g \cdot Y_{all} - (g - 1) Y(j) \quad j = 1, 2, 3, \dots, g \quad (1)$$

The pseudo-values can be used to set approximate confidence limits, using Students'  $t$ , as if they were independent, identically distributed random variables. The jackknife statistics,  $Y^*$ , and an estimate of its variances,  $S^{*2}$  are given by:

$$Y^* = \frac{1}{g} (Y_1^* + \dots + Y_g^*) \quad (2)$$

$$S^2 = \frac{\sum_j Y_j^{*2} - \frac{1}{g} (\sum_j Y_j^*)^2}{g - 1} \quad (3)$$

$$S^{*2} = S^2/g \quad (4)$$

#### Methodology

Data were obtained from 49 newly diagnosed essential hypertensives, prescribed diuretics. Patients from a Midwest and two Southwest medical centers were interviewed three to five weeks after they received their prescriptions. This questionnaire obtained patients' self reports of compliance, experience with side effects, attitude toward the prescribing physician and demographics. In addition, a 20-item measure of comprehension of drug regimen and a health locus of control measure (Walleston *et al.*, 1977) were administered. The health locus of control measure has intuitive appeal but is as yet untested in compliance studies. Finally, a behavioral measure of compliance, namely deviation from the hypothetical compliant refill date, was obtained from pharmacy archives.

#### Results

The data were analyzed with two linear models. First, using variables suggested by the compliance literature, a discriminant function was calculated for classifying compliant and noncompliant patients as they were so indicated in self reports.<sup>1</sup> While the model was able to classify nearly 75% of the cases correctly, none of the variables had significant discriminatory power.

The second model was a regression analysis of the same variables on refill deviance. The dependent measure, number of days noncompliant per 100, was modified to prevent outliers. That is, a ceiling of 35 days was implemented since patients who discontinued early in the data collection period had a head start toward infinity deviance. The results of this analysis showed that only the health locus of control measure had sig-

<sup>1</sup>Variables in the discriminant function included: self report of side effects (0-1), comprehension, health locus of control, attitude toward physician, sex, some college (0-1), and children at home (0-1).

nificant ( $p < .1$ ) variance reduction capability. Patients scored as internals were more compliant than their external counterparts.

Still, this apparently significant factor of compliance may be attributable to overdetermination or extensive searching of the data set. It represents a heretofore unnoticed factor of compliance that could add to (the confusion in) the literature. Before this insight receives acclaim it must meet the test of validation.

#### Validation via the Jackknife

The relationship between compliance and health locus of control was tested for stability by the approach previously outlined. More specifically, the data were divided into 41 groups<sup>2</sup> of size  $h = 1$  and the regression model was estimated 41 times. The  $n$  regression runs allowed calculations of pseudo-values and jackknife statistics for the regression coefficient of the health locus of control variable.

The results of the jackknife analysis are shown in Table 1. The coefficients' jackknife statistics, .045, divided by its corresponding standard deviation, .033, provides a  $t$ -statistic of 1.38. This is significant at the .2 level only. Hence, the health locus of control measure does not offer significant variance reduction in compliance.

TABLE 1

Jackknife Statistics for the Health Locus of Control Regressed on Compliance

$Y_{(j)}$	$\beta$ Coefficient	$Y^*_j$
$Y_{all}$	.049280	N/A
$Y_{(1)}$	.048513	.07996
$Y_{(2)}$	.046854	.14632
$Y_{(3)}$	.044529	.23932
$Y_{(4)}$	.048174	.09352
.	.	.
.	.	.
.	.	.
.	.	.
$Y_{(41)}$	.054984	-.17888
.	$\bar{x} = .04560$	
.	$s^* = .03303$	
	$t$ -statistic = 1.38 ( $p < .2$ )	

<sup>2</sup>Only 41 measures of refill deviance were obtained as some patients still had had no need to refill their initial prescriptions.

#### Conclusion

The preceding pages have served two worthy ends. First, the review of the compliance literature has provided a springboard for other consumer researchers to enter the area. Product liability issues and the social costs of noncompliance are just two factors sparking the interests of pharmaceutical and policy makers in this field of consumer behavior (Weintraub, 1975).

Future efforts will be more meaningful if conclusions can be validated. Toward this end, the administrative and cost factors which restrict sample sizes in the compliance setting also render traditional validation methods infeasible. This paper, however, outlines and demonstrates a jackknife procedure for more efficiently validating small sample conclusions.

Although the usefulness of the jackknife was demonstrated only on simple regression coefficients, its applications are extensive. Straightforward extensions to discriminant functions (Crask and Perreault, 1977) and even analysis of variance (Dwyer, 1978) are possible.

#### References

C. B. Bakker and C. R. Dightman, "Psychological Factors in Fertility Control," Fertility and Sterility, 15 (1964), 559-67.

W. L. Ball, "Improving Patient Compliance with Therapeutic Regimens: Hamilton Symposium Examines the Problems and Solutions," Canadian Medical Association Journal, III (August 3, 1974), 268 +.

M. H. Becker, R. H. Drachman, and J. P. Kirscht, "Predicting Mothers Compliance with Pediatric Medical Regimens," Medical Care, 81 (1972), 843-54.

D. Berry, A. Ross, and K. Deuschle, "Tuberculosis Patients Treated at Home," American Review of Respiratory Diseases, 88 (1963), 769-72.

Barry Blackwell, "The Drug Defaulter," Clinical Pharmacology and Therapeutics, 13, 6 (November - December, 1972), 841-848.

Barry Blackwell, "Patient Compliance," The New England Journal of Medicine, 289 (August 2, 1973), 249-53.

James R. Boyd, Tim R. Covington, Walter F. Stanaszek and R. Timothy Coussons, "Drug Defaulting. Part I; Determinants of Compliance," American Journal of Hospital Pharmacy, 31 (April 1974), 362-67.

James R. Boyd, Tim R. Covington, Walter F. Stanaszek and R. Timothy Coussons, "Drug Defaulting. Part II: Analysis of Non-Compliance Patterns," American Journal of Hospital Pharmacy, 31 (May 1974), 486-91.

R. H. Brook, F. A. Appel, and C. Avery, "Effectiveness of Inpatient Followup Care," New England Journal of Medicine, 285 (1971), 1509-14.

H. S. Caron, and H. P. Roth, "Patients' Cooperation with a Medical Regimen: Difficulties in Identifying the Non-Cooperator," Journal of the American Medical Association, 203 (1968), 922-26.

James C. Clinite and Hugh F. Kabat, "Prescribed Drugs ... Errors During Self Administration," Journal of the American Pharmaceutical Association, NS9, 9 (September 1969), 430-52.

- Melvin R. Crask and William D. Perreault, Jr., "Validation of Discriminant Analysis in Marketing Research," Journal of Marketing Research, 14 (February 1977), 60-68.
- F. Robert Dwyer, "An Experimental Study of Patient Use of Supplementary Drug Information," unpublished Ph.D. thesis, University of Minnesota, Minneapolis, 1978.
- Michael Etgar, "Channel Environment and Channel Leadership," Journal of Marketing Research, 14 (February 1977), 69-76.
- V. Francis, B. M. Korsch, and M. J. Morris, "Gaps in Doctor-Patient Communications," New England Journal of Medicine, 280 (1969), 535-40.
- M. S. Gatley, "To Be Taken as Directed," Journal of the Royal College of General Practitioners, 16 (1968), 39-44.
- Ernest Gurwich and Sister Emmanuel, "A Comprehensive Health Care Program," Journal of the American Pharmaceutical Association, NS14, 2 (February 1974) 71-79.
- E. H. Hare and D. C. Wilcox, "Do Psychiatric In-Patients Take Their Pills," British Journal of Psychiatry, 113 (November 1967), 1435-39.
- Daniel A. Hussar, "Patient Noncompliance," Journal of the American Pharmaceutical Association, NS15, 4 (April 1975), 183-90 +.
- G. S. M. Kellaway and Ellen McCrae, "Non-Compliance and Errors of Drug Administration in Patients Discharged from Acute medical Wards," New Zealand Medical Journal, 81 (June 11, 1975), 508-12.
- J. A. Leistyna and J. C. MacAuley, "Therapy of Streptococcal Infections: Do Pediatric Patients Receive Prescribed Oral Medication?" American Journal of Diseases and Children, 111 (1966).
- R. S. Lipman, K. Richels, E. H. Ehlenhuth, L. C. Park and S. Fischer, "Neurotics Who Fail to Take Their Drugs," British Journal of Psychiatry, III (September 1965), 1043-49.
- Edward E. Madden, Jr., "Evaluation of Outpatient Pharmacy Patient Counseling," Journal of the American Pharmaceutical Association, NS13, 8 (August 1973), 437-43.
- Bernadine Malahy, "The Effect of Instruction and Labeling on the Number of Medication Errors Made by Patients at Home," American Journal of Hospital Pharmacy, 23 (June 1966), 283-92.
- Rupert Miller, "The Jackknife -- A Review," Biometrika, 61 (1974), 1-15.
- Frederick Mosteller and John W. Tukey, Data Analysis and Regression. Reading, Massachusetts: Addison-Wesley, 1977.
- Richard L. Oliver and Philip K. Berger, "Testing Competing Models of Consumer Decision Making in the preventive Health Care Marketplace," in H. Keith Hunt, ed., Advances in Consumer Research, 5 (Ann Arbor: Association for Consumer Research, 1978), 277-82.
- A. M. W. Porter, "Drug Defaulting in a General Practice," British Medical Journal, 1 (January 25, 1969), 218-22.
- M. H. Quenouille, "Notes on Bias in Estimation," Biometrika, 43 (1956), 353-60.
- I. M. Rosenstock, "Why People Use Health Services," Milbank Memorial Fund Quarterly, 44 (July 1966), 94-127.
- D. L. Sackett and R. B. Haynes, Compliance with Therapeutic Regimens. (Baltimore: John Hopkins University Press, 1976).
- D. L. Sackett, R. B. Haynes, E. S. Gibson, B. C. Hackett, D. W. Roberts, D. W. Taylor, A. L. Johnson, "Randomized Clinical Trial of Strategies for Improving Compliance in Primary Hypertension," Lancet, 1 (1975), 1205-1207.
- D. Schwartz, M. Wang, I. Leitz, and M. E. W. Gross, "Medication Errors Made by Chronically Ill Patients," American Journal of Public Health, 52 (November 1962), 2015-19.
- Thomas R. Sharpe, "Potential Effects on the Patient," Drug Information Journal, 11 (January 1977), 58-64.
- Thomas R. Sharpe and Robert L. Mikeal, "Patient Compliance with Prescription Medication Regimens," Journal of the American Pharmaceutical Association, NS15, 4 (April 1975), 191-92 +.
- Jagdish Sheth, "What is Multivariate Analysis," in Jagdish Sheth, ed., Multivariate Methods for Market and Survey Research, Chicago: American Marketing Association, 1977.
- John W. Tukey, "Bias and Confidence in Not-Quite Large Samples," Annals of Mathematical Statistics, 29 (1958), 614.
- Barbara S. Wallston, Kenneth A. Wallston, Gordon D. Kaplan, and Shirley A. Maides, "Development and Validation of the Health Locus of Control (HLC) Scale," Nashville, Tennessee: Vanderbilt University, School of Nursing, 1975 (mimeo).
- Michael Weintraub, "Promoting Patient Compliance: Role of Health Professionals, Government, and the Pharmaceutical Industry," New York State Journal of Medicine (October 1975), 2263-66.



Albert R. Wildt, University of Georgia

## Introduction

Applications of rigorous research methodologies and statistical estimation techniques to problems of social consequence by consumer researchers are on the increase, as they probably should be. Often research of this type is intended to provide policy-makers with information to aid in the setting of public policy. In these cases it might be argued that researchers have a special obligation to provide accurate and valid information due to: (1) the importance, in terms of the degree of possible impact, of policy decisions, (2) the probable unfamiliarity of policy-makers with the techniques employed, as well as their possible inability to evaluate the resulting information, and (3) the likely impact of poorly conducted research on the receptiveness of policy-makers to this type of input for future decisions.

The discussion that follows considers three papers, all of which involve statistical estimation in the context of empirical research dealing with social and/or public policy issues. Barnes and Bourgeois (1978) consider the social problem of alcohol consumption in a study "...designed to explore primarily the question of whether increased governmental regulation of advertising might be expected to contribute to a reduction in per capita consumption of beverage alcohol products." Their research examines aggregate data (annual data for ten Canadian provinces) in an attempt to identify variables which have significant impact on annual per capita alcohol consumption during the period 1966 to 1973. Reizenstein and Barnaby (1978) examine the energy issue with the intent of providing insight to policy-makers. They employ a mail survey and estimate consumer response, in terms of quantity consumed, to hypothetical increases in the price of gasoline. Further, they group respondents according to similarities in response patterns into three market segments and identify these segments in terms of selected AIO measures, media and personal information source utilization variables and demographic measures. This is done in an effort to determine if the segments are "...sufficiently unique to warrant the formulation of market-oriented conservation strategies and tactics for specific groups." And Dwyer (1978) examines and summarizes research in the area of drug compliance (failure of patients to take medications as directed), mentions two sample-related problems with research in this area (small sample sizes and non-representative samples due to the nature of the sampling procedures employed) and indicates that in nonexperimental studies, such as those in the drug compliance area, where exploratory analysis procedures are employed it is very possible for otherwise spurious results to be interpreted as significant. Further, Dwyer suggests the use of the jackknife procedure to handle this latter problem and provides an illustration within the drug compliance area.

Two approaches come to mind for organizing the discussion of these papers. They are: (1) focus the discussion on the unique role of consumer research and statistical estimation relative to the "real" world of public policy setting, and (2) address selected statistical problems in estimation as they relate to consumer research and these three papers. Beckwith's (1978) paper in this same volume takes the former approach while the latter approach is pursued in this paper.

Each of the three papers considers estimation in one context or another. Barnes and Bourgeois are concerned with estimating the interrelationships among aggregate "social" variables, Reizenstein and Barnaby, among other things, are estimating consumer response through self-report intention measures, and Dwyer addresses the problem of the validation of estimates obtained when exploratory analysis methods are applied to samples of limited size. In this discussion paper, these specific estimation issues are considered under the general topic areas of: selection of appropriate models for estimation, measurement and estimation, and estimation using exploratory analysis procedures. This approach allows for both a brief general discussion of a set of issues in estimation and inference that relate to consumer research in general, and the detailed consideration of selected aspects of the three papers under discussion.

## Selection of Appropriate Models for Estimation

The model employed in any statistical analysis should be selected on the basis of its applicability to the specific problem under investigation. Unfortunately in some instances, the familiarity, or lack thereof, of the researcher with particular analysis models is the determining factor. This is not to imply that every researcher need be an expert on all analysis techniques, but to point out that it is important to be aware of the limitations of techniques so that inappropriate applications are avoided.

In many situations, especially when dealing with observational data, some doubt exists concerning the "correct" or appropriate statistical model. In instances where any doubt, no matter how small, exists, the researcher should, at a minimum, carefully examine the assumptions of the statistical model before it is employed and empirically test<sup>1</sup> those same assumptions after estimation. In some situations there are statistical model comparison procedures available for distinguishing among alternate models and, where applicable, these should be used.

## The Problem of Alcohol Consumption

In the Barnes and Bourgeois study dealing with alcohol consumption, the selection of an appropriate statistical model is clearly dependent upon the researcher's perception of the problem. Three modeling issues relating to this study are discussed below: (1) the appropriate degree of aggregation, (2) model specification relative to expected impact of variables, and (3) special considerations when estimating aggregate models.

Aggregate vs. Individual Models. The way in which the decision-maker frames the problem should dictate the analysis approach taken in any type of applied research. The alcohol problem is often described in relation to the heavy or chronic drinker. If this is viewed to be the appropriate focus, then one may argue that the researcher should investigate the heavy drinker and not

1 The testing of a model should also include considerations relating to the consistency of the estimated model with relevant behavioral or economic theory.

average consumption. Conversely, the problem may be considered in relation to young drinkers, in which case the focus of the research should be toward this particular segment. In either case, the level of aggregation employed in the present study could be questioned.

The authors do consider the aggregation issue and briefly address it by relating average alcohol consumption to the consumption of heavy drinkers. They point out that available data indicates that the distribution of alcohol consumption is approximately lognormal. Assuming the stability of this distribution over time, they contend that a shift in the parameter of central tendency of this distribution, average consumption, will be accompanied by a corresponding shift in the consumption of the heavy drinker. While this may or may not be a valid assumption, a possibly more plausible explanation for the use of aggregate data is the cost and difficulty of acquiring individual-level data as opposed to the availability of aggregate data.

Expected Impact. It might be effectively argued that the impact of most contemplated governmental regulation concerning this problem, especially that dealing with advertising, would be such that immediate, dramatic effects are not to be expected. That is, the timing of the impact, if any, might be some time in the future. For example, policy actions may result in discouraging present non-problem drinkers from becoming problem drinkers. Issues such as this are important and have serious implications on model specification.

Special Considerations for Aggregate Economic (or Social) Data. Much of consumer research is conducted with individual-level data. But, for certain problems, the use of aggregate data represents a desirable alternative. However, when using aggregate social or economic data it is important that the researcher adjust the model and estimation procedures to the unique characteristics of that data. A considerable amount of investigation into appropriate analysis methods when using aggregate economic data has been done by researchers in the area of econometrics. Econometrics is a special field of economics that deals with the measurement of relationships among economic (and social) variables. Goldberger (1964, p. 1) points out that, "although econometric theory draws heavily on the mainstream of mathematical statistics, it has a distinctive flavor which is attributable to characteristic features of economics. One feature is that as a rule observations on economic phenomena are not obtained by controlled experiments; consequently special methods for the analysis of nonexperimental data have to be devised. Another feature is that there is a rich body of theory of economic behavior; consequently special methods are devised to take advantage of this."

A number of issues frequently encountered in empirical econometrics are applicable to consumer research. Here, a limited number of those issues, which apply to the Barnes and Bourgeois study, are briefly discussed.

Many economic problems involve observations on cross-sections of time-series. Often in these instances there are unique characteristics of the cross-sections which impact directly on the dependent measure or possibly moderate the influence of the independent measures on the dependent variable(s). Also, observations from adjacent time periods often tend to be related to one another. To address these problems, special estimation procedures have been developed for use with time-series data from multiple cross-sections. (The interested reader may refer to Maddala [1977] for a brief discussion of this issue.)

In regard to the Barnes and Bourgeois study, it should be pointed out that the number of useable observations is exceedingly small (approximately 40) and the explicit

consideration of cross-sectional parameters in the estimation model would have severely reduced the degrees of freedom, thus causing other problems. However, the questions of cross-sectional differences and temporal relationships could still have been considered by the careful examination of the model which was estimated (e.g., residual analysis).

Also, when multiple dependent measures are employed the question of the interrelationships among these dependent measures arises. These interrelationships can assume any of a number of forms. For example, the multiple dependent measures may be functionally related, say a linear combination of each other, such as total alcohol consumption equaling the weighted sum of the consumption of beer, wine and spirits. (For a discussion of the impact of relationships such as these on modeling and estimation refer to Koehler and Wildt [1978] and McGuire and Weiss [1977].) In other applications the relationship may be indirect resulting in correlations among the dependent measures due to variables and relationships not explicitly included in the model (see Zellner, [1963] for a discussion of estimation procedures in this case). In any event, relationships such as these usually call for simultaneous, rather than separate modeling and estimation procedures.

The implication of using ordinary regression analysis (OLS) in situations which are characterized by the more complex models described above may be severe and are worth mention. At best, OLS will yield unbiased but inefficient estimators. And to complicate matters the estimated standard errors of those estimated coefficients will be biased, rendering inferences based on OLS estimates tenuous. At worst, the model will be misspecified and biased or even meaningless coefficient estimates will be obtained. Therefore, it is in the best interest of the researcher to carefully consider these issues.

#### Basis for Statistical Inference

An issue closely related to the selection of an analysis model is the basis use for statistical inference when multiple independent variables are considered. The options available to the researcher include simple, partial and sequential approaches. The simple approach considers each variable separately and ignores possible interrelationships with other variables. Reizenstein and Barnaby adopt this approach when they use a series of univariate F-tests to determine which variables best describe potential market segments. The partial approach considers only the unique aspects of the selected variable after accounting for all other variables. The regression models used by Barnes and Bourgeois, and Dwyer are examples of this approach. The sequential approach requires the researcher to develop a hierarchy within the set of variables and considers the unique aspects of the selected variable after accounting for only the variables above that variable in the hierarchy. Each approach may be useful and has a place in consumer research, although the partial approach is more common with observational studies, such as the three presently under consideration.

Specific analysis procedures often incorporate (either directly or due to available computational procedures) one of these three bases of inference. Therefore, the researcher must make a conscious decision concerning the basis of statistical inference when he selects an analysis model.

#### Measurement and Estimation

Empirical research frequently necessitates the development of operational definitions of variables and the construction of measurement scales based on these

definitions. In some cases these measurements are less than precise and might be better thought of as estimates (representations) of an underlying phenomenon or concept. Also, the measurement methods used impact on the nature of the measurements and, therefore, on the accuracy and validity of subsequent analyses.

#### Gasoline Consumption and Purchase Intentions

A critical measurement and estimation issue in the research reported by Reizenstein and Barnaby is the estimation of consumer response to hypothetical or anticipated changes in the levels of decision variables. This is a difficult problem and, if constrained to a survey research design, self-report intention measures appear to be a reasonable approach. However, a major concern here is the accuracy of consumers' response to questions concerning their behavioral intentions with respect to some future or hypothetical action. This is especially true if the action relates to conditions with which the respondent is unfamiliar or where existing or perceived social pressures favor a particular response. In the latter case, respondents who are more socially aware and concerned may yield more biased responses. The possibility of this problem effecting the accuracy and validity of the Reizenstein and Barnaby research is very real and a closer look at the measures employed in this study is warranted.

Unfortunately, the measurement instrument used in this study is not provided, however, under almost any conceivable measurement scale one would doubt the ability of respondents to give accurate responses.<sup>2</sup> Support for the accuracy of the intentions measure might have been provided through the measurement of variables relating to commitments by the consumer which might correlate with intentions. No such measures were reported. The large number of consumers indicating a 7-12 gallons per week reduction in consumption might provide some indication of the accuracy of the intentions measure. Assuming that the respondents were responding to reductions in personal (i.e., nonbusiness) consumption of gasoline, how realistic are these results? What would a 7-12 gallons per week reduction mean to you? According to the Statistical Abstract of the United States (Table No. 995, page 597) average annual U.S. gasoline consumption per car (includes taxicabs and motorcycles) in 1974 was 676 gallons, on 13 gallons per week. Given this average consumption, how realistic is it to expect 10% to 30% of the drivers to reduce consumption 7-12 gallons per week? On carefully considering the problem, it is probably reasonable to expect only small changes in consumption, especially in the short-run, in reaction to a price increase. And lastly, the aggregate data presented in Figure A (Reizenstein and Barnaby, 1978) indicates some minor data inconsistencies which should be carefully examined on an intrarespondent level.

While it is easy to point out the existence of the problem, it is a much more difficult matter to find a solution. Some consumer researchers may suggest that

---

2 The authors defend the "potential" validity of their intentions measure by arguing that the scale used has many properties of a purchase probability scale and refer to Juster (1966) for support concerning the validity of the purchase probability scale. However, it should be pointed out that Juster's work involved consumer durables and considered the probability of purchase, not the number of units purchased. Also, according to Green (1977, p. 106) "...the Bureau of the Census discontinued the Consumer Buying Expectations Survey [which employed such intentions measures] in 1973 because it was concluded that the data it provided were only 'marginal' useful."

experimental methods might yield more accurate indications of purchase intentions. However, given time and cost constraints, it is extremely difficult, if not impossible, to develop an accurate measure of purchase intention, especially within the context of survey research; though careful questionnaire design may go a long way in improving the quality of the intentions measure. It should also be noted that the research reported here is probably not very sensitive to inaccuracies in the prediction of the exact amount of decrease in gasoline consumption. The approach used by the authors requires only that respondents be grouped according to their relative propensity to decrease consumption. Therefore it would appear that only ordinal properties are required of the intentions measure.

#### Choice of Operational Definition of Variables

A somewhat different problem arises in connection with the Barnes and Bourgeois study and relates to the measurement of monetary quantities. When using time series data, it is usually advisable to measure monetary variables in terms of constant dollars rather than current dollars. When measured in current dollars, monetary variables often exhibit positive correlations to each other because of a common movement due to inflationary conditions. In the case at hand it not clear whether price and per capita income are measured in current or constant dollars. But if measured in current dollars, the positive trend of these variables would be accentuated because of inflation. This could well lead to positive relationships between these variables and consumption in alcohol, possibly due to the coincidental changes in the value of the monetary unit. Also, unnecessary multicollinearity may be introduced which would result in a decrease in the precision of the estimation and a general lessening in the ability to separate the impact of one variable from that of another.

#### Estimation Using Exploratory Analysis Methods

##### General Considerations

A common occurrence in consumer research is the use of exploratory data analysis methods, such as stepwise regression analysis. Often these analysis procedures are used with data sets of limited size and little or nothing is done in the way of validating results. The typical disclaimer is that the number of observations is too small to allow the use of a holdout sample for validation. The three studies considered here are examples. Barnes and Bourgeois conduct four stepwise regression analyses with nine independent variables in their estimation of the impact of selected variables on per capita consumption of alcohol. Reizenstein and Barnaby conduct 46 univariate F-tests in an effort to uncover variables which might differentiate three market segments and, for the two sets of results reported, 6 and 7 of the 46 were statistically significant at  $\alpha = .05$  and 2 and 3 at  $\alpha = .01$ . This is slightly more than would be expected by chance. In both of these instances no validation of results was undertaken. Dwyer employed stepwise regression analysis with seven independent variables and found one significant at  $\alpha = .10$ . Dwyer did attempt to validate his results, a point which will be discussed later.

As mentioned by Dwyer (1978), because of their tendency to capitalize on spurious correlations in the data, data searching techniques often uncover "significant" relationships that are, in fact, random. Therefore, results from these types of analyses should be carefully evaluated before any conclusions are reached.

Based on observation, one might argue that there are too many exploratory studies with results that are unvalidated. Two possible solutions to this problem may be worth considering. First, reduce the number of exploratory studies. Through the development of a scientific tradition whereby research hypotheses are generated based on the careful consideration of available theory, rather than the data, the reliance on exploratory studies could be reduced. This approach appears to be necessary when the researcher is limited in the number of observations available for analysis. In these cases the researcher must rely heavily on existing theory to formulate the model, and the iterative approach to model building and testing has the serious drawback of using the data to formulate the model and then using the same data to test the model. Procedures such as these raise serious doubts concerning the validity of any conclusions which may be reached. The second alternative is for researchers to take the necessary steps to validate results obtained from exploratory analyses, either through the collection of sufficient data to allow for a holdout sample or by some other means.

### The Jackknife

Dwyer (1978) suggests the use of the jackknife procedure as a method of validation which requires no additional data collection and illustrates the method within the context of regression analysis. At first glance the use of the jackknife as a substitute for a holdout sample for validation is very appealing. However, before accepting the procedure we should examine available evidence. Unfortunately, Dwyer offers no evidence supporting the use of the jackknife procedure for this purpose, therefore, other references (Quenouille 1956, Gray and Schucany 1972, Miller 1974, Miller 1974, and Mosteller and Tuckey 1977) were consulted. According to Gary and Schucany (1972, p. v.) "the jackknife is a general method for reducing the bias in an estimator and for obtaining a measure of the variance of the resulting estimator by sample reuse. Thus the result of the procedure is usually a nearly unbiased estimator and an associated approximate confidence interval." Mosteller and Tuckey (1977), in discussing the direct assessment of the variance of estimators, suggest the usefulness of the jackknife in those cases where insufficient observations are available to utilize the standard method of equivalent subsamples.<sup>3</sup> In this same regard, the jackknife is especially useful in complex multivariate problems where little or no theory exists to give exact tests of significance on coefficients, e.g., discriminant analysis.

Now, let's examine the applicability of the jackknife to regression analysis in light of the two above-mentioned benefits. First consider the issue of unbiasedness. The jackknife does not always yield an unbiased estimator but usually reduces the bias of a given estimator. If the regression model is correctly specified both the OLS and the jackknife give unbiased estimators. Under a misspecified model the OLS estimator will often be biased, but in this case the jackknife is also likely to be biased and since both estimators rely on the same incorrect model, neither may be of much use. In any event, research concerning the properties of jackknife estimators under conditions of misspecified regression models appears to be very sparse.

3 If the issue of concern to Dwyer was the direct assessment of variability, Mosteller and Tuckey (1977) would probably suggest dividing the sample into subsamples (after all there are 41 observations and only a single independent variable) in such a way as to measure the variability attributable to other relevant variables, such as hospital and length of time since receiving prescription. Alternately, these variables could have been included in the model.

Second, consider the jackknife as a method to obtain a direct measure of variance of an estimate. The usefulness of this is obvious in those complex cases where little or no theory exists concerning the distribution of sample estimates (refer to Crask and Perreault [1977] for such an application to discriminant analysis). But is it needed in the case of the regression model? If the regression model is correctly specified, it can be shown that the jackknife will have higher variance than the unbiased OLS estimator, and it is generally expected that the estimated variance of the jackknife will be greater than that of the OLS estimate. If the structure of the model is correctly specified except for the error term (e.g., non-normal or non-spherical disturbances), the direct assessment of variability could be very useful. However, this may not be necessary since there are methods available for detecting misspecifications of this type which should be used by the careful researcher. In the case of a structurally misspecified model, the jackknife estimator is also based on the misspecified model, therefore, depending on the nature of the misspecification, the usefulness of the jackknife is questionable. (Note: this may be a problem for future research.)

The discussion above presented arguments relating to the use of the jackknife in regression analysis to reduce bias or to obtain a measure of variability. Next, in a more general context, let's examine the question of whether the jackknife provides the same information as would an additional validation sample. The use of a validation sample is similar to randomly dividing the population into  $j$  samples. Even when there is no association between (among) variables in the population, spurious association may be observed in some of the  $j$  samples. The probability of this occurring is derived from sampling theory. When using exploratory data analysis procedures which capitalize on chance, the likelihood of uncovering significant "non-results" may be large, but the chance of confirming these results on the second (validation) sample is controlled by the  $\alpha$ -level selected. However, given a single sample with observed (though spurious) association, the further subdivision of that sample will not remove the association even though statistical tests on subsamples may fail to reject the null hypothesis due to either sampling variation and/or degree of freedom problems, or the use of less efficient estimation methods. The specific data observations comprising the sample are such that the relationship is present, and even the jackknife cannot change this. Since the jackknife is based on these same observations, one is forced to ask the question: Why should the jackknife do any better? It does appear that the jackknife will usually result in a more conservative estimate (i.e., higher variance), but if that is what one wants, why not select a smaller  $\alpha$ ? In summary, I would not be comfortable in recommending the jackknife as a substitute for a holdout sample given the evidence considered.

### Concluding Remarks

A number of interesting points are raised by the authors in these papers. But, in the opinion of this reader, there are two very important issues of a general nature addressed by the authors that deserve special mention. First, as pointed out by Barnes and Bourgeois for the problem of alcohol consumption and by Reizenstein and Barnaby in the case of energy, there exists the need for careful scientific investigation in the evaluation of contemplated public policy actions. An second, as Dwyer points out, researchers need to consider the problems inherent in the application of exploratory analysis methods to samples of limited size and the validation of results obtained under these conditions. This latter issue is one which everyone knows about but it appears to be the exception when somebody does something about it.

## References

- James G. Barnes and Jacques C. Bourgeois, "Estimating the Effects of Advertising: Application to a Social Problem," Proceedings of the Association for Consumer Research, 1978.
- Neil E. Beckwith, "Fiction, Folklore, Findings and Facts: A Discussion of Marketing Research in the Public Interest Forum," Proceedings of the Association for Consumer Research, 1978.
- Melvin R. Crask and William D. Perreault, Jr., "Validation of Discriminant Analysis in Marketing Research," Journal of Marketing Research, 16 (1977), 60-8.
- F. Robert Dwyer, "Drug Compliance and the Neglected Concern for Validity," Proceedings of the Association for Consumer Research, 1978.
- H. L. Gray and W. R. Schucany, The Generalized Jackknife Statistic (New York: Marcel Dekker, Inc., 1972).
- Arthur S. Goldberger, Econometric Theory (New York: John Wiley & Sons, 1964).
- Paul E. Green and Donald S. Tull, Research for Marketing Decisions (Englewood Cliffs: Prentice-Hall, Inc., fourth edition, 1978).
- Thomas F. Juster, Consumer Buying Intentions and Purchasing Probability (New York: National Bureau of Economic Research, 1966).
- Gary Koehler and Albert R. Wildt, "Characterization and Estimation of Admissible Logically Consistent Parameters for Constrained Linear Models," Research Report No. 78-6, Industrial and Systems Engineering Department, University of Florida, Gainesville, Florida, May 1978.
- G. S. Maddala, Econometrics (New York: McGraw-Hill Book Company, 1977).
- Timothy W. McGuire and Doyle L. Weiss, "Logically Consistent Market Share Models II," Journal of Marketing Research 13 (1976), 296-302.
- Rupert G. Miller, "An Unbalanced Jackknife," Annals of Statistics, 2 (1974), 880-91.
- \_\_\_\_\_, "The Jackknife - A Review," Biometrika, 61 (1974), 1-15.
- Frderick Mosteller and John W. Tuckey, Data Analysis and Regression (Reading, Massachusetts: Addison-Wesley Publishing Company, 1977).
- M. H. Quenouille, "Notes on Bias in Estimation," Biometrika, 43 (1956), 353-60.
- Richard C. Reizenstein and David J. Barnaby, "Assessing the Potential Effects of Differential Price Increases on Gasoline Usage," Proceedings of the Association for Consumer Research, 1978.
- U.S. Department of Commerce, Bureau of the Census, Statistical Abstract of the United States: 1976 (Washington, D.C.: U.S. Government Printing Office, 97th edition, 1976).
- Arnold Zellner, "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias," Journal of the Americal Statistical Association, 57 (1962), 348-68.

CHANGES IN CONSUMER PERCEPTIONS:  
THE IMPACT OF TESTING CONDITIONS ON PERCEPTIONS OF BRANDED PRODUCTS

James McCullough, University of Arizona  
Charlene S. Martinsen, University of Washington  
Linda Sceurman (Student), University of Washington

Abstract

Perceptions of products as measured by multidimensional scaling of perceived differences appear to be insensitive to changes in the stimulus set than had been expected. Selection of elements for presentation to subjects in this type of study does not appear to be a critical source of experimenter bias.

Measurement of Perceptual Change

Although the use of multidimensional scaling (MDS) techniques to examine perceptions of sets of products has become of fairly common approach in studies of consumer behavior (Green, 1975), a major difficulty in the use of MDS has been interpretation of the perceptual spaces derived by the technique. The particular problem of dimensional interpretation has been discussed in detail (Shepard, 1974). Studies of simple stimuli indicates dimensional structure is related to the attributes of the stimuli presented (McCullough, Martinsen and Moinpour, 1978). If this is true, then it seems likely that alteration of the set of stimuli presented would change the underlying dimensional structure of the perceptual space derived by MDS.

The method of presentation of stimuli has been shown to influence both the perceptions of the objects (Allison and Uhl, 1964) and the weightings assigned to the perceptual dimensions (Rao, 1972). These researchers have clearly demonstrated that product perceptions and subsequent perceptual spaces are highly sensitive to changes in explicit information presented concerning stimuli at the time of evaluation. Situational effects on perceptual spaces have been discussed by McCullough and Martinsen (1976), particularly when comparing brand evaluations with sensory evaluation of actual products. An excellent taxonomy of the relationship between perceptions and underlying structure can be found in Garner (1976). These concerns have stimulated this study which examines changes in expressed perception induced by changes in the elements of a stimulus set and by alternative methods of stimulus presentation.

Multidimensional scaling (MDS) appears to be a particularly good technique for examining changes in product perception induced by changes in the method of presentation of the products. Moinpour, McCullough and MacLaclin (1976) have demonstrated the usefulness of MDS in measuring longitudinal changes and have shown that measurements of perceptual spaces can be replicated with a high degree of reliability. That study examined the impacts of messages about products when the stimuli (brandnames) were presented in a consistent fashion. This study examines the changes in product perceptions induced by alteration of the presentation format and by modification of the stimulus set.

Changes in Brand and Product Perceptions

It has been shown that the underlying structure of the perceptual space for a set of product can be related to the characteristics of the elements of the stimulus set (McCullough, Martinsen, and Moinpour, 1978). It seems reasonable from this finding that by changing elements in

a set the perceived structure could be altered. It seems likely that these changes would occur: 1) if dissimilar elements were introduced into the stimulus set, or 2) if the set of attributes of the product presented to the subject is changed, since both represent significant alternation of the stimulus set.

The changes measured by MDS could result from three changes in perception by the consumer: 1) The attribution of the dimensions of the space to the stimuli may be changed; that is, the subject may apply the discrimination criteria differently to existing stimuli; 2) the subject may use new discriminating criteria or may reduce the number of criteria used; or 3) the subject may change the relative weighting or importance of the discriminating criteria. If the structure of the perceptual space is more stable than expected and is not altered by changes in the stimulus set, the spatial relationships between the elements of the stimulus set should remain relatively unchanged.

To attempt to measure the effect on the structure of the perceptual space of changes in stimuli three types of change were examined: 1) alteration of the set attributes as perceived by the subjects; 2) substitution of a single stimulus similar in form to existing stimuli in the stimulus set 3) substitution of several stimuli of different form in the stimulus set. It was believed that these treatments would provide a reasonable range of conditions under which changes might be induced.

Method

Subjects and Training

Predominately female groups of approximately 20 subjects between the ages of 18 and 35 were chosen to participate in the various phases of this study. The subjects were familiar with sensory evaluation techniques but had no prior experience with MDS procedures. They were told that they were taking part in a test of a new marketing research tool, that they would be asked to evaluate a set of soft drink brands and that they were to use any criteria they felt were appropriate. Subjects then participated in a tasting session in which they were asked to taste and evaluate a set of soft drink samples. Subjects were given no information concerning the brands presented for sensory evaluation until the scoring had been completed.

Stimulus Presentation

Previous studies had employed sets of commonly consumed and readily available soft drinks selected to represent the variety found in the marketplace. The single house brand in the stimulus set was selected as the element to be substituted since it was likely to have less impact on perceptual structure than the national brands. For the multiple substitution treatment, diet analogs of brands were substituted into the stimulus set while attempting to maintain the analogous product clusters which had been previously found in the standard set.

All products used in sensory evaluation were purchased

in 12 ounce cans at local supermarkets and were refrigerated prior to serving. One ounce samples of soft drinks were presented to the subjects in small glasses for sensory evaluation. The soft drink brands used in the study are shown in Table 1.

Subjects completed a questionnaire requiring similarity-dissimilarity ratings on nine point scales for all possible pairs of products or brand names of the products under study. The pairs of stimuli were evaluated in randomized order, and each subject responded to the same order of pairs. Most individuals required 15-30 minutes to make the required pairwise evaluations.

Proximity ratings provided by each subject submitted to the INDSCAL individual differences scaling algorithm (Carroll and Chang, 1970). The resulting stimuli spaces were compared using the C-MATCH algorithm (Cliff, 1966). Individual saliences obtained from INDSCAL were compared across groups using the MANOVA algorithm where appropriate (Cooley and Lohnes, 1971).

.95 had been reported between brand perceptions and sensory evaluations. In that study McCullough and Martinsen (1976) had reported significant changes in dimensional salience despite high correlations between spatial maps. Based on those findings it was suggested that verbal surrogates should not be employed in marketing research applications aimed at measuring actual product perceptions because perceptions in the two presentation modes likely would differ. However, they were unable to demonstrate any significant difference in the perceptual spaces between brand perceptions and sensory evaluations of soft drinks, in spite of significant differences in dimensional salience between the two modes. Their conclusions only served to raise questions concerning the methodology.

In the present study, the perceptual maps were so clearly different that the MANOVA test for differences in salience was unnecessary and seemed inappropriate since the actual underlying dimensions had obviously changed. This strengthens the previous findings that suggested

TABLE 1  
EXPERIMENTAL STIMULI INCLUDED IN SOFT DRINK PERCEPTUAL SETS

<u>Standard Set</u>	<u>Single Stimulus Substitution</u>	<u>Multi-Stimulus Substitution</u>
Pepsi Cola	Pepsi Cola	Pepsi Cola
Tab <sup>a</sup>	Tab <sup>a</sup>	Tab <sup>a</sup>
Seven-Up	Seven-Up	Seven-Up
Sprite	Sprite	Sprite
Dr. Pepper	Dr. Pepper	Dr. Pepper
Shasta Lemon-Lime	Shasta Lemon-Lime	Diet Pepsi <sup>a</sup>
Coca Cola	Coca Cola	Diet Seven-Up <sup>a</sup>
Cragmont Lemon-Lime <sup>b</sup>	Cragmont Cola <sup>b</sup>	Diet Dr. Pepper <sup>a</sup>

<sup>a</sup>diet soft drink

<sup>b</sup>house brand soft drink

#### Results and Discussion

Subject dissimilarity ratings were submitted to the INDSCAL algorithm to identify the brand position in the perceptual space most common to the subjects. The solution obtained for the standard set is shown in Figure 1. Perceptual spaces obtained for the various treatment conditions, were compared using the C-MATCH algorithm on the coordinates of all stimuli common to the spaces being compared. That is, only coordinates of stimuli present in both spaces were used in making spatial comparisons. The degree of the shift in the eight stimulus spaces induced by substituting a single element was measured by comparing the positions of the remaining seven stimuli, while comparison of the standard set with the high diet set was based on comparisons of the five common elements. These analyses were performed for both sensory evaluation of products and for evaluation of brand names. The results of these analyses are shown in Tables 2 and 3.

#### Sensory Evaluation

For common non-diet soft drink brands, a C-MATCH goodness of fit of .96 and a distance vector correlation of

verbal representations of actual soft drink products should be employed with caution in marketing research situations.

TABLE 2  
COMPARISON OF PERCEPTUAL SPACES USING C-MATCH  
GOODNESS OF FIT FOR INDSCAL SOLUTIONS FOR SENSORY  
EVALUATIONS

	Correlation (Goodness of Fit)
Standard with multiple substitution	.41 <sup>a</sup>
Multiple substitution with multiple substitution under blind conditions	.95 <sup>b</sup>
Brand evaluations with sensory evaluations	.57 <sup>b</sup>

<sup>a</sup>based on five stimuli

<sup>b</sup>based on eight stimuli

TABLE 3  
COMPARISON OF PERCEPTUAL SPACES USING C-MATCH  
GOODNESS OF FIT FOR INDSICAL SOLUTIONS FOR  
BRAND NAME EVALUATIONS

	Standard	Single Substitution
Single Substitution	.98 <sup>a</sup>	--
Multiple Substitution	.97 <sup>b</sup>	.99 <sup>b</sup>

<sup>a</sup>based on seven stimulus coordinates

<sup>b</sup>based on five stimulus coordinates

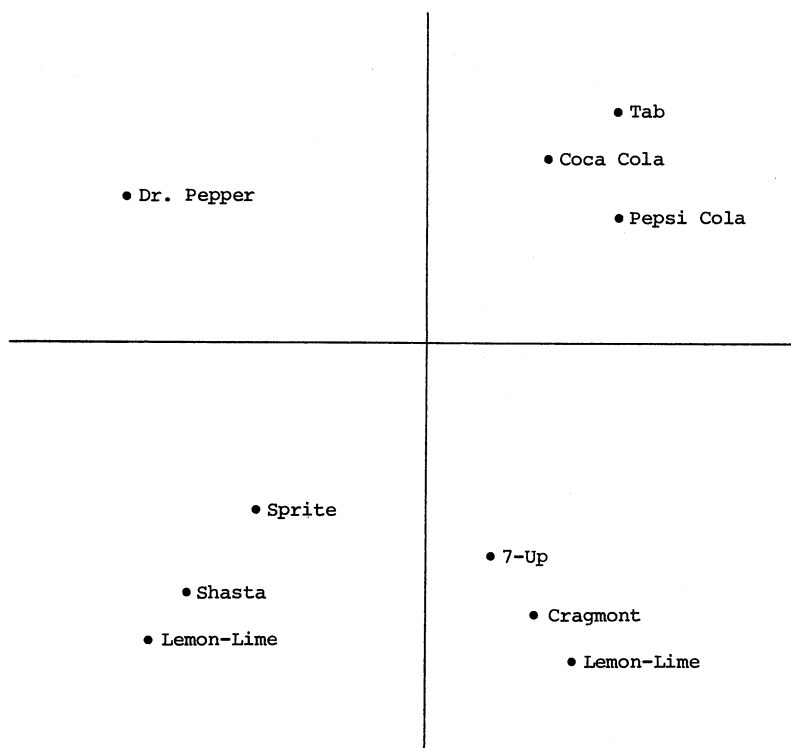
relative positions of the seven constant elements. The result shown in Table 3 indicate a clear consistency between the two spaces and no significant alteration of the space was found.

Substitution of three diet brands for three conventional soft drinks was expected to induce a shift in spatial dimensionality to include a diet/non-diet dimension altering the relative positions of the five constant elements which included Tab, also a diet drink. The goodness of fit of .99 and the vector correlation of .98 indicate that such a shift did not occur.

Examination of the mean subject correlations with the INDSICAL perceptual map for the various spaces indicates a possible, slight alteration of the underlying perceptual structure as elements are altered. The single brand substitution improved the mean square correlation from .76 to .83 while the substitution of diet brands

FIGURE 1

A PLOT OF THE INDSICAL STIMULI CONFIGURATION FOR THE STANDARD SOFT DRINK SET



In attempting to explain the reasons for the significant differences noted in this study, it was believed that color represented a significant discriminating variable in the sensory evaluation of the soft drink samples that might be influencing perceptions. To test the impact of color, subjects were presented samples in covered containers which hid the color of the samples. There was no significant difference between the configurations and the goodness of fit was .95. These results are shown in Table 2. Apparently the differences in sensory perceptions are more complex than had been hypothesized.

#### Brand Evaluations

Substitution of a single element in the stimulus set was not expected to produce any significant change in the perceptual maps. The goodness of fit was .98 and the correlation of distance vectors was .95 based on the

reduced the correlation to .73. These shifts are slight and not statistically significant indicative of a tendency toward higher dimensionality.

When the perceptual spaces are not structurally dissimilar, as occurred between sensory evaluations and brand perceptions of diet soft drinks, a MANOVA test for salience dissimilarity would seem to be appropriate. The utilization of the dimensional structure by the subjects as indicated by dimensional salience may change even when the perceptual maps appear identical, and examination of possible variation in dimensional salience using MANOVA can provide additional insight concerning product perceptions. In this case, however, the MANOVA analysis failed to indicate any significant differences between the groups tested. Group mean saliences for the groups tested are shown in Table 4. The only significant differences are between the standard sets in



sensory and brand evaluations as had been previously reported (McCullough and Martinsen, 1976).

#### Comparison of Sensory and Brand Evaluations

Earlier research had shown that configurations derived from brand evaluations were quite similar to those derived from sensory evaluation. This was not the case

TABLE 4

GROUP MEAN SALIENCES FROM INDSCAL ANALYSIS  
FOR GROUPS EXHIBITING SIMILAR PERCEPTUAL SPACES

Group	Dimension	
	I	II
Standard Sensory Evaluation	.65	.33
Standard Brand Evaluation	.82	.28
Single-Stimulus Substitution Brand Evaluation	.81	.29
Multi-Stimulus Substitution Brand Evaluation	.78	.28

for the multiple substitution set. The goodness of fit for these two perceptual spaces was only .57. This may indicate that for conventional soft drinks brand perceptions are much closer to the actual consumption experience than for diet drinks with which the consumers may have had less experience. This is clearly only a hypothesis for which additional experimental support is necessary.

#### Conclusions

The results of this study would seem to indicate that the selection of elements for inclusion in stimulus sets for evaluation by MDS may not be a critical source of experimenter bias, and that subjects are able to maintain a stable perceptual space even when elements are substituted. Since one would expect that the selection of particular stimuli would play a large role in the determination of the underlying dimensions of the perceptual space developed by individual subjects, this raises a question about the basis for dimensional determination.

One common technique identifying the axis of MDS solutions involves use of maximum "r" regressions against attribute ratings for individual stimuli. It is generally assumed, however, that the subject is free to select any criteria deemed appropriate in developing the perceptual space, and the resulting space contains the actual perceived distances between stimuli in that space. If alteration of the stimulus set fails to produce changes in the perceived relationships between unaltered elements, as occurred in this experiment when additional diet drinks were substituted, then the dimensions of the space may not be linear axes in the space but might more correctly be viewed as areas or domains as Shepard (1974) suggested. This question should be explored in more detail.

At a practical level, it would appear that alterations of a stimulus set which do not change the basic perceived character of the set can be incorporated. The experiment involving a single element change clearly indicates that is true. The failure of the multiple substitution to change the space supports this, since extensive changes can apparently be made and the perceived

relationships between unaltered stimuli will be retained.

This study has not attempted to thoroughly catalog changes which can or cannot be made. The effects of large and small changes have been examined and characterized for a single product category. Further experimentation in this area is necessary to adequately characterize effects for stimuli which are not so familiar as soft drinks.

#### References

- R.I. Allison and K.P. Uhl, "Influence of Beer Brand Identification on Taste Perception." Journal of Marketing Research, 1964, 1, 36-39.
- J.D. Carroll and J.J. Chang, "Analysis of Individual Differences in Multidimensional Scaling Via an N-Way Generalization of Eckart-Young Decomposition." Psychometrika, 1970, 35, 283-319.
- N. Cliff, "Orthogonal Rotation to Congruence." Psychometrika, 1966, 31, 31-42.
- W.W. Cooley and P.R. Lohnes, Multivariate Data Analysis. New York: John Wiley and Sons, 1971, 223-242.
- W.R. Garner, "Interaction of Stimulus Dimensions in Concept and Choice Processes." Cognitive Psychology, 1976, 8, 98-123.
- P.E. Green, "Marketing Applications of MDS: Assessment and Outlook." Journal of Marketing, 1975, 24, 410-420.
- J.M. McCullough and C.S. Martinsen, "Differences in Stimuli Evoked Responses Between Physical Stimuli and Verbal Surrogates." Proceedings of the American Marketing Association Educators' Conference, 1976, 45-48.
- J.M. McCullough, C.S. Martinsen and R. Moinpour, "Application of Multidimensional Scaling to the Analysis of Sensory Evaluations of Stimuli with Known Attribute Structures." Journal of Applied Psychology (in press).
- R. Moinpour, J.M. McCullough and D.L. Machachlan, "Time Changes in Perception: A Longitudinal Application of Multidimensional Scaling." Journal of Marketing Research, 1976, 13, 245-53.
- V.R. Rao, "Changes in Explicit Information and Brand Perceptions." Journal of Marketing Research, 1972, 9, 209-213.
- R.N. Shepard, R.N. "Representation of Structure in Similarity Data: Problems and Prospects." Psychometrika, 1974, 39, 373-421.

COVARIANCE BIAS OF THURSTONE CASE V SCALING  
AS APPLIED TO CONSUMER PREFERENCES AND PURCHASE INTENTIONS

Joel Huber, Duke University  
Murphy A. Sewall, the University of Connecticut<sup>1</sup>

Abstract

A possible source of bias in Thurstone Case V scaling as applied to buyer preference and intention to purchase data is examined. An empirical test for the existence of spatially related covariance bias is proposed and applied to data from two field studies.

Introduction

In spite of its use by marketing professionals (Green and Tull, 1978; Sewall, 1978a), there are good reasons to believe that the usual assumptions of Thurstone Case V scaling (1927) are systematically violated by the data collected in many buyer behavior studies. In a typical marketing context, Thurstone Case V analysis is applied across subjects to preference judgments on competing brands. The scale then provides a model of the proportion of respondents that will prefer one brand over another. Thus, as a model, it reduces the  $n(n-1)/2$  of original proportions on pairs to a more readily manageable scale with 'n' values.

The underlying concept, dubbed a "discriminal process," is simple and intuitively compelling. Without getting into details that may be found in Edwards (1957), Torgerson (1958), or Bock and Jones (1968), the process assumes that each evaluation of an object is drawn from a distribution with fixed mean and variance. When a pair of items is compared, the item with the higher momentary value is chosen. In order to estimate the hypothesized mean affective values that make up the scale, it is necessary to make the assumption, among others, that pairs of objects have momentary values with equal covariances. It is this assumption that is consistently violated in many applications of Case V across subjects.

Briefly, covariance bias arises if there is heterogeneity of preferences across subjects that is related to a common perception of the similarities between the objects under study. Under such conditions, the result is a distortion in the predicted pairwise probabilities that can be directly related to the degree of pairwise dissimilarity.

Conceptual Analysis

There are two forms of bias which may occur as a result of violating the Case V assumption of equality of covariances. The first form derives from cases where individuals' affective values are random deviations from the group mean. In this model, pairs that share an item in common will be correlated for a respondent since a high probability of choosing that item on one pair implies a higher probability of choosing that item when it is paired with other objects. Bock and Jones (1968, p. 144ff) provide a relatively complex, but usable, procedure to adjust for this source of heterogeneity of covariances.

The second form of bias currently is not analytically tractable. This source of bias arises when the variation in preferences across subjects is not random, as in

the case above, but instead is based upon a spatial model. That is, respondents who tend to like one object also tend to like other objects which are spatially similar. This is not a new idea; it is the necessary condition for most attribute models and has severe implications with respect to several probabilistic choice models. To quote Bock and Jones (1968, p. 133):

"Indeed the principle of independence from irrelevant alternatives has the same effect as constant correlation of discriminational processes for all pairs of stimuli. For, if the correlations are constant and the discriminational dispersions are equal, the difference process for pairs of objects, not sharing an object in common are uncorrelated. This implies that the conditional probability of a subject's choice between two objects, given his choice between any other two objects, is equal to the unconditional probability. It is doubtful that the principle is strictly true in practice, since personal preferences for given objects usually extend to similar objects. (A person who likes turnips will probably also like rutabaga and may even have a taste for parsnips.) Nevertheless, the assumptions of Luce's model and Thurstone's Case V may appear to be well enough approximated in many applications to allow reasonably accurate predictions of choice."

It is this last assertion that may be questionable in a consumer behavior context. There is good reason to believe that systematic distortions in choice due to perceptual similarities are not insignificant. Indeed, Coombs (1964, ch. 8) bases an entire joint-space scaling on the assumption that correlations in affective values across subjects enable one to recover the latent dimensionality of the space.<sup>2</sup>

Furthermore, any model of preference based on attributes implies that the correlation between objects will depend upon attribute similarity. That is, a subject's liking for one object implies a greater probability of liking similar items. This means that similar items will tend to be positively correlated while dissimilar items will tend to be negatively correlated.

The Thurstone Case V assumes that all pairs of objects have identical covariances. If, contrary to this assumption, similar pairs are more highly correlated, a consistent bias will result, moving predictions for these pairs closer to 0.5. This bias can be clearly seen by examining an extreme case.

Suppose two similar items have a (covert) correlation of nearly 1.0. In such a situation, one item would almost always dominate the other because its momentary value would increase whenever the other's did. Case V, however, assumes that all pairs have the same covariance. Thus, the predicted probability for this pair would be

<sup>1</sup>Also Senior Research Associate, Cambridge Marketing Group of New York City.

<sup>2</sup>It should be noted that Coombs based his derivation of the joint space of subjects and objects on the correlation of affective values between subjects. But since the joint space is symmetric between subjects and objects, it could be identically defined from the correlation of affective values between objects, as above.

biased closer to 0.5. Similar logic applies to dissimilar items, except the bias is expected to be away from 0.5 and towards either 0.0 or 1.0.

#### A Test of This Distortion

Modeling this distortion is complicated by the fact that the direction of the predicted bias reverses when the predicted probability is below, rather than above, 0.5. One way of eliminating this problem is to recode the predicted proportions on all pairs to be greater than 0.5. Because, by definition, the predicted probability of choosing 'j' over 'k' is the complement of the probability of choosing 'k' over 'j' ( $p_{jk} = 1 - p_{kj}$ ), it is a simple matter to make all predicted probabilities greater than 0.5 by reordering the subscripts.

To avoid the problem of scale effects wrought by probabilities, all probabilities are transformed to standard normal deviates (z-scores) by the inverse of the cumulative normal probability distribution. Assuming that the relationship between the bias and some measure of dissimilarity between pairs is approximately linear in the z-scores, the strength of the relationship can be tested by

$$z_{jk} - \hat{z}_{jk} = a + b \cdot dj_k, \quad (1)$$

where

$$\hat{z}_{jk} > 0.0$$

If bias is systematically in the predicted direction, the coefficient of the dissimilarity between two objects ( $dj_k$ ) should be negative. This sign occurs because a highly dissimilar pair will tend to have a negative (covert) covariance, and the predicted value will be biased upward, creating a negative residual term. Conversely, a highly similar pair ( $dj_k \approx 0$ ) will result in a  $\hat{z}_{jk}$  that is biased downward so that residuals will be positive.

The spatially related covariance distortion in Thurstone scales also is expected to be an increasing function of the affective heterogeneity of the underlying subjects. That is, bias will be greater if individuals have widely different preferences for the objects under study.

#### A Field Test

A Marketing application of Thurstone Case V scaling was reported at the 1977 Association for Consumer Research meeting (Sewall, 1978a). The scaling reported in that study is based on derived paired comparisons calculated from five-point intention to purchase ratings. It was pointed out at the meeting that, since derived paired comparisons suppress intransitive pairs, the usual statistical test for the adequacy of Thurstone Case V (Mosteller, 1951) produces an exaggerated estimate of accuracy. Hence, the test proposed in this paper appears to be a more appropriate method of evaluating scales based upon derived paired comparisons.

Sewall (1978b) also has developed a procedure for dividing a population into intention to purchase or preference segments based upon the rank order correlations between individual subjects and a group scale. Because the procedure requires that all subjects have a significant rank order correlation with the Thurstone scale for the segment, one would expect much greater affective homogeneity for subjects within a segment than for the population as a whole. Thus, covariance distortion is expected to be much greater for a scale based on all subjects than for scales for each segment.

#### Comparison Data

Intention to purchase ratings for two different product types were collected for this study. Both data sets are from mall shopper intercept interviews taken in four major metropolitan areas in different sections of the United States. Five-point intention to purchase ratings on 22 bed linen (sheet) designs were collected in the Spring of 1977. Data on 29 drapery designs were gathered in the Summer of 1977. All interviewing was done by paid professional interviewers.

The data from the two surveys provide contrasting prior expectations about the severity of covariance distortion for total population scales. The two largest segments found in the drapery study are of nearly equal size and quite divergent in their purchase intentions. The responses to the sheet survey are dominated by a large segment containing about sixty percent of the subjects.

Table 1 contains the correlation matrices between the segment Thurstone scales for each survey. Because the correlations between the scales for the two large segment scales (numbered 1 and 2) in the drapery survey are significantly negatively correlated, greater affective heterogeneity is implied in this population than for the sheet survey respondents. The one large segment in the sheet survey (numbered 1) has a Thurstone scale that is less divergent from the scales for the two smaller segments. Hence, the population scale for the sheet survey should be dominated by the purchase intentions of the largest segment, and less covariance distortion should be found in the group scale for the subjects responding to the sheet survey than for respondents to the drapery survey.

TABLE 1

Correlations Between Preference Segments' Derived Scales

Segment	Drapes			
	1	2	3	4
1	1.0	-.726	.724	.281
2		1.0	-.839	.212
3			1.0	.060
4				1.0

Segment	Sheets		
	1	2	3
1	1.0	.113	.466
2		1.0	-.447
3			1.0

Another measure of the relative homogeneity of subjects upon which a Thurstone scale is based is the average rank order correlation between the subjects' ratings and the scale values. Scales for groups with lower average subject correlations are expected to produce greater evidence of covariance distortion than scales for groups with higher average subject correlations.

#### Dissimilarities Data

The independent variable for each data set is based upon dissimilarities in physical attributes of the products. Several obvious characteristics of the designs (for example, weave, pattern, lining, and variety of color for the draperies) were coded as dummy variables (either the characteristic is present in the design or it is not). The matrix of designs by attributes was subjected to principal components analysis and dissimilarities between pairs is operationalized as the Euclidian distance between designs in reduced space. For the drapery designs, the first two components (containing 79.4 percent of the variance in the four attributes) are used as a basis for

dissimilarities.

Ten attributes were identified for the 22 sheet designs. The first three principal components (containing 60.8 percent of the attribute variance) are used as a basis for calculating dissimilarities.

#### Least-Squares Analysis

Values of the dependent and independent variables (equation 1) were plotted to determine if a nonlinear transform of the data might be appropriate. The plots indicate the relationship between bias and dissimilarity is approximately linear.

Table 2 contains correlation results and calculations of average subject to Thurstone scale rank correlations. Significant evidence of spatially related covariance distortion is indicated. The results are generally consistent with expectations about the behavior of such distortion. In all instances, the direction of the bias is as predicted. Distortion is greatest in the group scale from the drapery survey. The dominance of segment 1 in the sheet survey is indicated by a higher average subject correlation for the total sample compared to the average subject correlation for the drapery survey respondents. Within surveys, distortion is greater for scales based on total samples than for scales based on segments of relatively homogeneous individuals. However, evidence of significant distortion appears in the segment scales, and the average subject correlation does not appear to be a general indicator of the severity of this distortion.

Table 2  
Relation Between Case V Residuals and Dissimilarities

Segment	Drapes		Average Subject Correlation
	Correlation	F(1,404)	
All	-.412	82.76	.148
1	-.245	25.78	.588
2	-.176	12.95	.575
3	-.224	21.39	.569
4	-.113	5.26	.455

Segment	Sheets		Average Subject Correlation
	Correlation	F(1,229)	
All	-.189	8.46	.342
1	-.114	3.01	.558
2	-.128	3.84	.575
3	-.019	0.08	.472

For simple regression, correlation equals the standardized regression coefficient.  
For the F-ratios,  $\alpha(p<0.05)$  is 3.84;  $\alpha(p<0.01)$  is 6.63.

#### Discussion

Under the assumptions of Thurstone Case V, the probability that a particular object will be chosen over another can be derived from the cumulative normal probability distribution. Although the logic is quite different and the estimation methods differ slightly, the estimated probabilities are virtually identical to models which assume the logistic function (Luce, 1964) or the arcsin transform distribution (Torgerson, 1958; Bock and Jones, 1964; David, 1963). Therefore, covariance bias applies

to these models as well.<sup>3</sup>

Subjective dissimilarities on objects may be more appropriate for detecting distortion than the objective dissimilarities used in this study. Consumers are likely to react to some attribute differences more strongly than others. It is possible that evidence of spatially related covariance distortions might be even stronger if subjective dissimilarity data were available.

From a marketing management perspective, the existence of covariance distortion on Thurstone Case V predictions for pairs will have little effect on the average affective values relevant to such questions as:

1. What types of designs are favored by various consumer segments?
2. What types of designs are unlikely to appeal to any segment of consumers?

Thus, although predictions of choice of one object over a dissimilar one may be exaggerated, the bias is unlikely to alter a conclusion about which type of dissimilar objects is favored.

If management wishes to choose between similar objects, distortion may cause a problem. Because predictions understate choice probabilities, the scale may not indicate the degree to which one object actually is favored over another, similar one.

It should be noted, however, that the covariance distortion found to exist in affective judgments is less likely to be a problem when applied to perceptual judgments; that is, judgments that one item is larger, creamier, or zettier than another. Perceptual judgments should be less susceptible to covariance bias because they are more homogeneous across subjects than preference judgments.

#### Conclusions

Spatially related covariance distortion can exist in Thurstone Case V scales derived in consumer preference or purchase intention studies. Significant distortion can occur in data which appears acceptable on the basis of an established test for the adequacy of Thurstone Case V scaling (Mosteller, 1951). The adjustment for covariance distortion requires separate data on inter-object dissimilarities, but is particularly appropriate for Thurstone scales of preference based upon derived paired comparison data aggregated across consumers.

Procedures that cluster respondents into subjects which express similar preference or purchase intention orderings appear to reduce the magnitude of this spatially related covariance distortion. Whether this distortion should concern management users of Thurstone scale information depends specifically on whether management is concerned with overall affective values or on the differences between individual items. In the latter case steps should be taken to reduce covariance distortion.

#### References

Darrell R. Bock and Lyle V. Jones, The Measurement and Prediction of Judgment and Choice (San Francisco:

<sup>3</sup>An analysis was run using arcsin transformed variables (Mosteller, 1951) instead of z-values with results which were not significantly different from those reported in Table 2.

Holden-Day, 1968).

C. H. Coombs, A Theory of Data (New York: Wiley, 1964).

H. A. David, The Method of Paired Comparisons (New York: Hafner, 1963).

A. L. Edwards, Techniques of Attitude Scale Construction (Englewood Cliffs, New Jersey: Prentice-Hall, 1957), 19-82.

P. E. Green and D. S. Tull, Research for Marketing Decisions, 4th edition (Englewood Cliffs, New Jersey: Prentice-Hall, 1978), 258-9.

R. D. Luce, Individual Choice Behavior (New York: Wiley, 1959).

Frederick Mosteller, "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.

Murphy A. Sewall, "Nonmetric Unidimensional Scaling of Consumer Preferences for Proposed Product Designs," in H. Keith Hunt, ed., Advances in Consumer Research, vol. 5 (Ann Arbor, Michigan: Association for Consumer Research, 1978), 22-5.

\_\_\_\_\_, "Market Segmentation Based on Consumer Ratings of Proposed Product Designs," Journal of Marketing Research (1978), forthcoming.

L. L. Thurstone, "A Law of Comparative Judgment," Psychological Review, 34 (1927), 278-86.

Warren S. Torgerson, Theory and Methods of Scaling (New York: Wiley, 1958).

## MARKET SHARES ESTIMATES BASED ON CONJOINT ANALYSIS OF CONCEPTS

James B. Wiley, University of Florida  
Robert Bushnell, Wayne State University

### Abstract

This paper describes a market share model for estimating the proportion of "first choices" that will be received by each of an arbitrary number of concepts. The model is particularly useful for estimating market shares of multiattribute alternatives (products) using conjoint analysis.

### Market Shares Estimates Based on Conjoint Analysis of Concepts

Conjoint analysis is a recent development in mathematical psychology that is concerned with measuring the effect of two or more variables on the ordering of a dependent variable.<sup>1</sup> "Conjoint", of conjoint analysis, refers to the notion that it may be possible to measure the relative effects of two or more variables even though their effects individually may be unmeasurable. Conjoint analysis differs from related procedures, such as dummy variable regression or ANOVA, in that only ordinal assumptions are made regarding the dependent variables. Consumer psychologists have found conjoint analysis very useful in the study of consumer preferences for multiattribute alternatives.

The task addressed is to model, fit, and if successful, to predict the choices among alternatives where each alternative is a concept described by a list of its possession of described levels of certain attributes. By a concept we mean a conceptualization of a real object as distinct from the object itself. The descriptive dimensions of the attribute and the descriptions of the levels of the attributes which concepts could possess are all prepared by the interrogator.

It should be noted that the term "attribute" has several interpretations in consumer research. First, the use of the term may suggest a physical characteristic of an object, such as the sweetness of a drink or the color of a car. Second, its usage may imply criteria for membership in a category, as, for example, "a 'strike' in baseball might be defined as either a pitch that the batter swings at and either misses (A) or hits in foul territory (B), or a pitch to the batter in the area above home plate (C), above the batter's knees (D), and below his armpits (E); that is,

$A \cup B \cup (C \cap D \cap E)$  strike". (Wyer, 1974, p. 20.)

Finally, it may be used as a way of classifying specific perceptual attributions about members of another class, such as a product. With this usage, "sportiness" is interpreted not as something cars have in common, but rather something that attributions regarding cars may have in common.

Whatever the usage, however, multiattribute alternatives, or concepts, may range in abstractions from verbal descriptions, through pictorial representations, to physical "mock-ups" (Green and Tull, 1978). Regardless of

the level of abstraction, however, in the mind of the interrogator each concept is regarded as a combination of attributes, each possessed at a specified level. The set of concepts presented to respondents, then, are designed to contain individual concepts or combinations of levels of attributes which are systematically varied from each other. Further, restricting the choice situation by controlling the stimuli assures that each concept is evaluated with respect to the same information. Ambiguous and equivocal cues are removed, and all respondents are thereby certain to have the same information at their disposal and no more so that the inferences made beyond this point will only have their origins in the data provided. (Hoffman, 1960.)

Of course, attributes are not the only things that influence individuals' preferences. As a matter of practice, however, those who use conjoint analysis generally divide the factors that influence consumer preferences into two types. Attributes of concepts which influence alternatives' relative "favorableness" (or utility) comprise the first type. The following assumptions are made regarding attributes:

- (1) The consumer can distinguish "levels" or states of attributes (representing specific attributions).
- (2) The consumer attaches part-utilities (or partworths) to these levels or states.
- (3) The overall utility (or favorableness) of a concept is a function of its part-utilities.
- (4) In any set of alternatives, the concept having the greatest utility will be most preferred.
- (5) Choice will be a function of preference.

Factors of the second type include characteristics or states that are not attributed to concepts, but which may influence preferences. Examples include the context in which this consumer will use the concept, the needs that are fulfilled through its use, institutional constraints (such as distribution channels), and policies of parties that can influence choice. Factors of this type can influence the outcome of conjoint analysis through two mechanisms. First, they can operate to exclude potential alternatives from consideration. Secondly, they can operate by influencing the way in which the consumer chooses amongst those concepts that he is presented. For example, the relative importance of an attribution may depend on the context of the choice.

Three sets of procedures are necessary to implement conjoint analysis (Johnson, 1974):

- (1) A technique of data collection requiring consumers to consider "trade-offs" among attributions,
- (2) A computational method which derives partworths by accounting as nearly as possible for each consumer's choice behavior, and
- (3) A model which allocates "market share" to concepts according to the relative utility of the competing concepts.

<sup>1</sup>Work on this paper was supported by ERDA grant "Future Transportation Systems of the Great Lakes Area: Energy and Economics". Contract E(11-1)-4136.

To date, the majority of work in the area has focused on the first two sets of procedures.

It is not sufficient, however, to account for the preferences of a single consumer. Market share is the result of many decisions made by many consumers. Each of these consumers may differ in the utilities ascribed to attribute levels and, hence, consumers may disagree in their preferences for specific alternatives. This paper presents a model for estimating the market share that will be captured by each of an arbitrary number of concepts when choices are made by a heterogeneous group of decision-makers.

### Conjoint Analysis

In order to infer the part-utility of attributions, conjoint analysis requires the assumption of a composition model which specifies the mechanism by which the attribute scores interact and are related. An additive model is frequently adopted. It is then assumed that a consumer bases his choices on the utility of alternatives and that utility is determined according to the following model (the additive conjoint analysis model):

$$U_{ij} = \sum_{k=1}^n f_{ik}(x_{jk}) \quad (i=1, \dots, n; j=1, \dots, q; k=1, \dots, m) \quad (1)$$

where,  $U_{ij}$  = the overall utility of alternative j for individual i.

$f_{ik}$  = a real valued function defining the utility on the kth attribute for the ith individual, and

$x_{jk}$  = the level on attribute k characterizing the jth alternative.

Note the similarity between model (1) and main effects analysis of variance in that in both cases an overall score is assumed to be the sum of a set of effects defined on attribute levels (Green and Wind, 1973).

If the hypothesized rule holds to a satisfactory degree of approximation, we consider the matrix U with elements  $U_{ij}$  defined as in (1):

$$U = \{U_{ij}, i=1, \dots, n; j=1, \dots, q\} \quad (2)$$

Each row of this matrix represents the ith individuals utility for each of the q alternative concepts. Each column represents the utility scores on the jth concept registered by the n respondents. Given this matrix, we require a procedure which will estimate for each concept the mean utility score  $\bar{U}_j$ , its standard deviation,  $\sigma_j$ , and the correlation of its utility scores with the utility scores of the other concepts. Pekelman and Sen (1977) discuss one approach for estimating these parameters. A weighted least squares procedure is also possible. Both procedures assume attributes are measured at the interval level.

### Estimating Market Share

We now turn to the problem of predicting market share, i.e., the proportion of times each concept will be chosen when decision-makers are given a choice of one out of n concepts. If  $U_i$  is the utility of the ith alternative, ( $i=1, \dots, n$ ) then  $U_i^*$  is chosen if  $U_i^* > U_i$  for all  $i \neq k$ , that is,  $U_i^* = \max \{U_i, i=1, n\}$ . If for simplicity, we consider initially the problem of predicting the proportion of choices allocated one concept among three, then for concepts  $C_i, C_j$ , and  $C_k$  we require the probability of the event (Bock and Jones, 1968):

$$(U_i > U_j) \cap (U_i > U_k) \quad (3)$$

If we designate the difference between utilities within the above parentheses as  $u_{ij} = U_i - U_j$  and  $u_{ik} = U_i - U_k$ , respectively, the required probability may be written as  $P[(u_{ij} \cap u_{ik}) > 0]$ . If in the population of consumers, the distribution of utilities assigned  $C_i, C_j$ , and  $C_k$  are assumed  $N(U_i, \sigma_i), N(U_j, \sigma_j)$ , and  $N(U_k, \sigma_k)$ , with  $\text{Cov}(U_i, U_j) = \rho_{ij}\sigma_i\sigma_j$ , etc., the joint distribution of  $u_{ij}$  and  $u_{ik}$  is bivariate normal. This distribution may be denoted as  $N(u_{ij}, u_{ik}, \sigma_{ij}, \sigma_{ik}, \rho_{ij, ik})$  where  $u_{ij}$  is the difference between the mean utility assigned  $C_i$  and the mean utility assigned  $C_j$ ,  $\sigma_{ij}$  is the standard deviation of the utility differences, and  $\rho_{ij, ik}$  is the correlation between the difference scores for  $(C_i, C_j)$ . The proportion of times  $C_i$  will be judged to exceed both  $C_j$  and  $C_k$  by a randomly selected consumer is given by the positive quadrant of the bivariate normal distribution.

Integration of the multivariate normal distribution when the number of variables exceeds two is extremely difficult. While such problems can be solved using numerical analysis, it generally is not practical to apply a multivariate normal market share model when the number of concepts exceeds three, since this would involve integration of a trivariate (or higher dimensioned) distribution of utility differences. Fortunately, if the correlations between utilities can be assumed equal to one-half, a useful approximation is available for cases involving four or more concepts.

### A Logistic Market Share Model

The bivariate logistic distribution in its reduced form is defined as:

$$F(x, y) = [1 + e^{-x} + e^{-y}]^{-1} \quad (4)$$

Gumbel (1961) has shown that the marginal expectations are zero and the marginal standard deviations are  $\pi/\sqrt{3}$ . Furthermore, the covariance between the marginal scores is  $\pi^2/6$  and hence the correlation coefficient has the fixed value one-half (1/2).

Bock and Jones (1968) suggest that the generalized multivariate distribution provides a convenient vehicle for calculating the proportion of first choices that will be received by each of the n alternatives. To use this function, it is only necessary to adjust the variates to fit the normal distribution by equating variances in the marginal distribution; that is, by using  $\pi/\sqrt{3}$  times the unit normal deviates as logistic deviates. When this is done in terms of the utility difference variables  $u_{ij}$  defined in the previous section (e.g.,  $u_{n1} = U_n - U_1$ ), the probability that  $C_n$  will be chosen first from among n concepts is approximated by:

$$P(U_n > U_1, U_2, \dots, U_n) = P(u_{n1} > 0) \quad (u_{n2} > 0) \quad \dots$$

$$u_{n, n-1} > 0 =$$

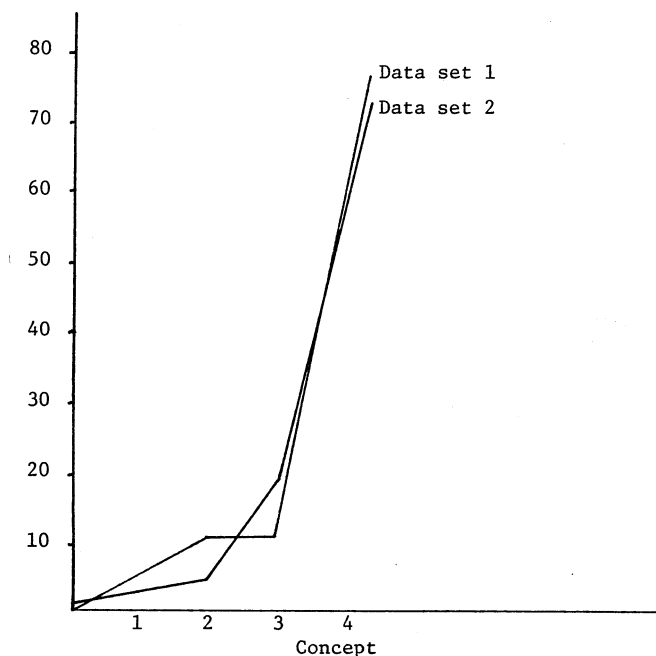
$$1 / (1 + e^{-\frac{\pi}{\sqrt{3}} \frac{u_{n1}}{\sigma_{n1}}} + e^{-\frac{\pi}{\sqrt{3}} \frac{u_{n2}}{\sigma_{n2}}} + \dots + e^{-\frac{\pi}{\sqrt{3}} \frac{u_{n, n-1}}{\sigma_{n, n-1}}}) \quad (5)$$

### Example

Figures 1, 2, and 3 illustrate the effect of various parameter changes on the estimated proportion of first choices. From these figures, it is evident that the function has at least three desirable characteristics. First, whatever the number of concepts or their utilities, the "market shares" assigned each will sum to one. Second, the greater the difference between the mean utilities of the concepts, the greater the difference between their market shares (assuming equal standard deviations).

This property is illustrated in Figure 1. Note that concepts "2" and "3" have equal mean utilities in the first data set and that they are assigned equal market shares.

FIGURE 1  
Effect of Utility Differences on Market Share Estimates

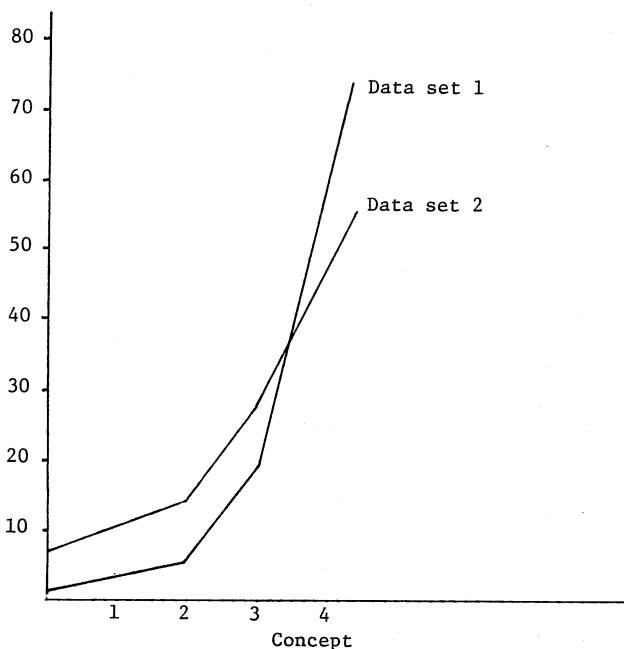


Concept	1	2	3	4
Data Set 1				
Mean Utility	1.0	2.5	2.5	4.0
Standard Deviation	1.0	1.0	1.0	1.0
Data Set 2				
Mean Utility	1.0	2.0	3.0	4.0
Standard Deviation	1.0	1.0	1.0	1.0

The third desirable property of the model is illustrated in Figure 2. This figure illustrates that the more the respondent population agrees about the utilities of the concepts (the smaller the variances), the more the model will discriminate between concepts on the basis of small utility differences. For example, the same utilities are assumed in both data sets; however, the standard deviations of the utilities in the second data set are four times the standard deviation in the first data set. Market share estimates in the small variance data set range from 1.5 percent to 72 percent while market share estimates in the large variance data set range from 7.5 percent to 54 percent even though both are based on the same mean utilities.

Figure 3 illustrates further the interesting market share dynamics that can occur as a consequence of interaction between the utility estimates and their variances. Here concepts having the same mean utilities have somewhat different market shares as a consequence of their differing standard deviations. In general, a concept will pick up market share relative to other concepts as the variance of its utility increases. This suggests that it is possible for a concept to have a smaller mean utility than alternative concepts and yet capture the largest "first choice" or market share.

FIGURE 2  
Effect of Uncertainty on Market Share Estimates



Concept	1	2	3	4
Data Set 1				
Mean Utility	1.0	2.0	3.0	4.0
Standard Deviation	1.0	1.0	1.0	1.0
Data Set 2				
Mean Utility	1.0	2.0	3.0	4.0
Standard Deviation	4.0	4.0	4.0	4.0

#### Comparison With Other Models

It is interesting to compare the behavioral assumptions of the logistic model with other models. (We follow Bock and Jones, 1968, in these matters.)

#### Relationship to Thurstone Case V

The psychologist Thurstone developed a number of models which link subjective properties of stimuli to observable choice proportions. The most restrictive model -- which also is the most practical -- incorporates the so-called Case V assumptions. Torgerson (1958, pp. 159-165) provides an excellent discussion of the Case V assumptions. The covariance between the difference scores can be represented as (Bock and Jones, 1968, p. 249):

$$\sigma_{ij}\sigma_{ik}\rho_{ij,ik} = E(U_i - U_j)(U_i - U_k) - E(U_i - U_k)E(U_j - U_k) = \sigma_i^2 - \sigma_i\sigma_j\rho_{ij} - \sigma_i\sigma_k\rho_{ik} + \sigma_i\sigma_k\rho_{jk}, \quad (6)$$

where

$$\sigma_{ij}^2 = \sigma_i^2 + \sigma_j^2 - 2\sigma_i\sigma_j\rho_{ij}, \quad \text{and} \quad (7)$$

$$\sigma_{ik}^2 = \sigma_i^2 + \sigma_k^2 - 2\sigma_i\sigma_k\rho_{ik}. \quad (8)$$

It is the term  $\rho_{ij,ik}$  -- which Bock and Jones call the



"Comparital correlation" -- that is assumed equal to 1/2 when using the multivariate logistic model. The central assumptions of Thurstone Case V are that  $\sigma_i^2 = \sigma_j^2$  for all  $i$  and  $j$  and that  $\rho_{ij} = \rho_{ik}$  for all  $i, j,$  and  $k$ . With these assumptions, equations (6), (7), and (8) can be written:

$$\sigma_{ij}\sigma_{ik}\rho_{ij,ik} = \sigma^2 - \sigma^2\rho - \sigma^2\rho + \sigma^2\rho = \sigma^2(1 - \rho - \rho + \rho) = \sigma^2(1 - \rho) \quad (6a)$$

$$\sigma_{ij} = \sqrt{\sigma^2 + \sigma^2 - 2\sigma^2\rho} = \sigma\sqrt{2(1 - \rho)} \quad (7a)$$

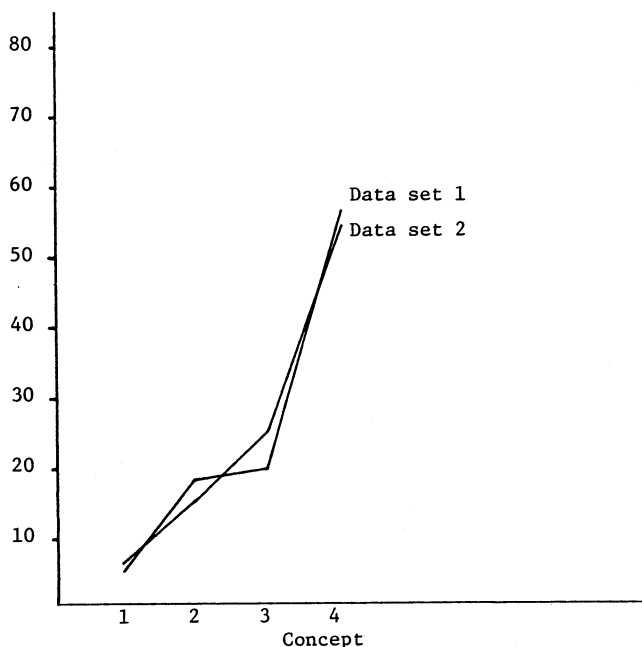
$$\sigma_{ik} = \sqrt{\sigma^2 + \sigma^2 - 2\sigma^2\rho} = \sigma\sqrt{2(1 - \rho)} \quad (8a)$$

Equation (6a) may then be solved for  $\rho_{ij,ik}$

$$\rho_{ij,ik} = \sigma^2(1 - \rho) / \sigma^2 \sqrt{2(1 - \rho)} \sqrt{2(1 - \rho)} = 1/2.$$

That is, if the Thurstone Case V assumptions hold, then the "comparital correlations" will equal one-half, consistent with the properties of the logistic model.

FIGURE 3  
Effect of Simultaneously Varying Parameters on Market Share Estimates



Concept	1	2	3	4
<b>Data Set 1</b>				
Mean Utility	1.0	2.5	2.5	4.0
Standard Deviation	1.0	2.0	3.0	4.0
<b>Data Set 2</b>				
Mean Utility	1.0	2.0	3.0	4.0
Standard Deviation	1.0	2.0	3.0	4.0

#### Relationship to the Bradley-Terry-Luce (B.T.L.) Model

Bradley and Terry (1952) and Luce (1959) have suggested a model to link probability of choice and preference. Their models assert that:

$$P(j) = V(j) / \sum_{k=1}^q V(k) \quad (9)$$

where  $P(j)$  = probability that individual will choose brand  $j$ ,  $V(j)$  = consumer's ratio scale preference for brand  $j$ ,  $k = 1, 1, \dots, j, \dots, q$ . Equation (9) may be re-written:

$$P(j) = \frac{1}{\frac{V(1)}{V(j)} + \frac{V(2)}{V(j)} + \dots + \frac{V(j)}{V(j)} + \dots + \frac{V(n)}{V(j)}} \quad (9a)$$

This equation may be converted to the logistic model by setting:

$$V(j)/V(k) = \exp [(U(j) - U(k)) / \sqrt{(\sigma_j^2 + \sigma_k^2 - 2\sigma_j\sigma_k\rho_{jk})}] = e^{u_{jk}/\sigma_{jk}} \quad (10)$$

Although the Bradley/Terry/Luce model (9) normally is applied to individual choice behavior, Bock and Jones (1968, sec. 6.5.1) provide an interpretation of model parameters that allows the model to be applied to populations of individuals. The B.T.L. model, therefore, can be considered a special case of (5).

Because B.T.L. model incorporates a probabilistic version of the axiom of independence of other alternatives, however, it cannot account for the effect that similarity between alternatives may have on choice. For example, suppose an individual is given a choice of a can of Green Giant peas ( $U = .3$ ), a can of Del Monte peas ( $U = .3$ ), and a can of Del Monte peas with a 10¢ rebate coupon ( $U = .4$ ). Assuming the alternatives have the same price, the B.T.L. model suggests the Del Monte can with the coupon has .40 probability of selection. If a Del Monte is selected, however, almost everybody would agree the version with the coupon would be selected. Note that whatever the valued characteristics of the regular can may be, they are shared with the coupon-bearing can. Consequently, a large and positive covariance between the utilities assigned these two alternatives is expected (10). This will result in a small standard deviation of difference scores and a large difference between "market shares" will be predicted for these two alternatives by the logistic market share model.

#### Conclusion

This paper discusses a logistic market share model that can be used in connection with conjoint analysis. The model appears to provide a practical vehicle for estimating market shares that would be captured by abstract alternatives, or concepts.

We make the following assumptions in developing the market share model:

- (1) decision-makers attach part-utilities (or part-worths) to concept characteristics,
- (2) the overall utility (or "favorableness") of a concept is a function of the sum of its part-utilities,
- (3) the most favorable concept will be selected.

Thus, individual choice is assumed to be deterministic. If all decision-makers agree on concept characteristics and the part-worths of these characteristics, then a single concept will receive all the choices, ceteris paribus.

However, all factors will not be equal. In particular, it is likely that decision-makers disagree on the part-

worths they assign characteristics. If decision-makers disagree on partworths, then a distribution will exist for the favorableness of each concept. Therefore, choice is stochastic in the aggregate.

The logistic market share model is somewhat more general than Thurstone Case V in that unequal standard deviations for utility difference estimates can be accommodated. It is somewhat more general than the B.T.L. model in that it is not necessary to assume that pairwise ratios of choice probabilities are independent from other alternatives.

It should be noted that McFadden (1970) has proposed a probabilistic individual choice model similar in formulation to the market share model outlined. It appears likely that the two formulations could be combined in a model having both individual and aggregate stochastic elements.

The logistic market share model as currently formulated does not allow external information to be incorporated into market share estimates. For example, a subset of the concepts presented individuals may describe actual alternatives for which market share data is available. A useful extension of the current formulation would enable the researcher to "calibrate" the model using prior data on a subset of concepts. Needless to say, it would also be desirable to include in the calibration process covariate information vis-a-vis the subset of alternatives, such as their distribution share, advertising share, promotion share, and so forth. Validation of the model against subsequent choices is of ultimate concern. A series of such studies is contemplated.

#### References

- R. D. Bock, and L. V. Jones, The Measurement and Prediction of Judgment and Choice, (San Francisco: Holden-Day, 1968).
- R. H. Bradley, and M. E. Terry, "Rank Analysis of Incomplete Block Designs, I, The Method of Paired Comparisons," Biometrika, 39(1952), 324-345.
- P. E. Green and Y. Wind, Multiattribute Decisions in Marketing: A Measurement Approach, (Hinsdale, Illinois: The Dryden Press, 1973).
- E. J. Gumbel, "Bivariate Logistic Distributions," Journal of the American Statistical Association, 56(1961), 335-349.
- P. J. Hoffman, "The Paramorphic Representation of Clinical Judgment," Psychological Bulletin, 47(1960), 116-131.
- R. M. Johnson, "Trade-Off Analysis of Consumer Values," Journal of Marketing Research, 11(May 1974), 121-127.
- R. D. Luce, Individual Choice Behavior, (New York: John Wiley & Sons, 1959).
- D. McFadden, "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka (ed.) Frontiers in Econometrics, (New York: Academic Press, 1970), 105-142.
- D. Pekelman and S. Sen, "Regression Versus Interpolation in Conjoint Analysis," Advances in Consumer Research IV, W. D. Perreault, Jr. (ed.), 29-34.
- W. S. Torgerson, Theory and Methods of Scaling, (New York: John Wiley & Sons, 1958).
- R. S. Wyer Jr., Cognitive Organization and Change: An Information Processing Approach, (New York: John Wiley & Sons, 1974).

Discussant's Comments

Robert C. Blattberg  
Graduate School of Business  
University of Chicago

COVARIANCE BIAS OF THURSTONE CASE V  
SCALING AS APPLIED TO CONSUMER PREFERENCES  
AND PURCHASE INTENTIONS

Joel Huber, Duke University  
Murphy A. Sewall, University of Connecticut

This article raises a number of important questions about the use of Thurstone scales based upon paired comparisons. The paper demonstrates that certain biases exist in the scale when the assumption of equal covariances is violated. This problem is definitely relevant to marketing and consumer researchers. My comments relate to the following: (1) the need for a theory of choice to demonstrate the key problems that they have found; (2) specific attempts to overcome this problem in the econometric literature.

A choice model is needed because the problem the authors are describing may have one of two causes: (1) correlation between products or implicitly between attributes of products or (2) a two-stage choice process. The second problem is irrelevance of independent attributes and is different from correlation. For example, suppose there are three soft drinks available, Pepsi, Coke, and Seven-up. If one were to use paired comparisons, they might find out that Pepsi and Seven-up were equally chosen by consumers or equally be preferred, Coke and Seven-up were equally preferred, and Pepsi and Coke were equally preferred. In terms of Thurstone's model, one would find out that these could be scaled at exactly the same point. If the consumer used a two-stage process in selecting brands, they may first select on the attribute cola-ness, in which case they may be equally likely to select Seven-up or a cola product. If they select a cola product, then they are equally likely to select Pepsi and Coke. If they select an "un-cola", i.e. Seven-up, there is 100% chance that they will buy Seven-up. Thus, the choice probabilities become .25, .25, .50 for Pepsi, Coke, and Seven-up respectively.

This is really the example that Bock and Jones are describing in the passage given in the article. Thurstone's underlying model does not consider a two-stage choice process. Therefore, the reason Thurstone's Case V does not model the data well may be because the consumer uses a different choice model than Thurstone assumed. Therefore, the problem discussed is really model misspecification bias rather than correlation bias.

Another key question that the authors have not spent much time on but allude to is the implications for designing new products. Again a mathematical formulation of the problem would greatly help the authors and the readers understand the implications for product policy. Without a formal analysis it is difficult to understand the implications of correlation bias or misspecification bias on product design.

Some attempts exist to overcome the correlation bias described. In the econometric literature, Hausman and Wise [1978] try to introduce correlations into a probit model, making it a multivariate probit model. They are trying to overcome the same problem this article considers. Hausman and Wise's solution requires very lengthy computer runs to do numerical integration of multivariate normal distributions. Therefore, it is not clear how easy it is to introduce different corre-

lations into Thurstone's Case V. The cost of solving the problem may be far greater than the biases associated with product selection or product design.

In summary, the authors have raised an important issue. The cause of the problem is ambiguous and the solution may be very costly. However, it offers a direction for some future research.

MARKET SHARES ESTIMATES BASED ON  
CONJOINT ANALYSIS OF CONCEPTS

James B. Wiley, University of Florida  
Robert Bushnell, Wayne State University

The question the authors study is: How does one convert the utilities from conjoint analysis into consumer choices? By making certain statistical assumptions about randomness of utilities, a Thurstone or Luce choice model is recommended for converting utilities into choice probabilities. There are a number of difficulties with this approach. First, the authors do not state the cause of the randomness in the model. Is the randomness in utilities due to heterogeneity in the population or variation across time in individuals' choices? This question is not carefully answered in the article but could result in a very different interpretation of the choice model.

Second, why does one need a complex choice model at all? Why not make the probability of choice simply equal to the fraction of first choices divided by the total number of observations drawn? This may be an accurate estimate of the probability of choosing the  $k^{\text{th}}$  configuration. This method is obviously very simple. The authors should state why this is inappropriate.

Third, a theory of choice is needed. Suppose the trade-off matrix is the one given in figure 1. The trade-off is between miles per gallon and size of the car. Suppose the consumer is willing to make the first choice, 40 miles per gallon in a large car; second choice, 30 miles in a large car; and third choice, 20 miles in a large car. According to the type of choice model or probabilities that are generated from Luce's model, one would find that it is possible to select a large car that gets 30 miles per gallon X percent of the time when a large car that gets 40 miles per gallon is available. In other words, because the utilities generate probabilistic choices, it is now possible to generate situations in which an individual is willing to choose a car that is less preferred to a car that is more preferred. Obviously, this decision makes very little sense since the consumer clearly prefers a car that gets 40 miles per gallon over 30 miles per gallon, holding everything equal. Yet applying the choice models recommended, this case can occur.

Thus, unless one is very careful about the choice model used, and the measurement of that choice model, cases which violate key choice axioms can be generated. My suspicion is that Luce's model or Thurstone's model is not consistent with the underlying concepts of trade-off analysis. The link between the underlying choice model and the measurement is not very carefully done in this paper or in other papers on conjoint analysis. Therefore, more work has to be done on the issue of linking choice models and the measurement models. I think this is a research direction the authors should seriously consider.

The authors have worried about a key question which has not been worried about very much in the conjoint analysis literature in marketing. They have obviously moved

this a step forward by considering some of the issues and raising a number of important questions. Further research in this area is strongly needed.

Figure 1

Trade-Off Matrix

		Size		
		Large	Medium	Small
Miles per gallon	40	1	4	7
	30	2	5	8
	20	3	6	9

Table entries are ranked preferences.

CHANGES IN CONSUMER PERCEPTIONS:  
THE IMPACT OF TESTING CONDITIONS ON  
PERCEPTIONS OF BRANDED PRODUCTS

James McCullough, University of Arizona  
Charlene S. Martinsen, University of Washington  
Linda Scurman (Student), University of Washington

The conclusions reached by the authors are important to researchers using Multidimensional Scaling (MDS). For MDS to be a reliable research tool, it is important that the spatial location of stimuli are added to the stimulus set. In order for this result to be widely accepted, two additional research steps need to be taken: (1) sample of "real" people used from the population at large, and (2) many more product categories analyzed. These two steps will insure that the results appearing in this paper are not merely a random occurrence due to a special product or a biased sample. The need to replicate the results across product categories is particularly important to avoid idiosyncratic results. If the result found in this paper is affirmed in these replications, it improves the validity of MDS.

James L. Ginter, Ohio State University

## Introduction

The common trait of the three papers in this session on "Methodological Developments" is that they all explore relevant questions in the use of scaling methods in marketing research. The first two papers [2,1] investigate potential problems with common applications of scaling methods, while the third paper [3] proposes a model for estimation of market share from scaling results. Each of the papers is based on a different method of measuring attitudinal constructs, with multidimensional scaling, Thurstone Case V, and conjoint analysis being used. The contribution of each of these efforts is explored through the following questions and comments.

"Changes in Consumer Perceptions: The Impact of Testing Condition on Perceptions of Branded Products"

The primary conclusion drawn from this research [2] is that measured perceptions of products, based on multidimensional scaling of perceived differences, tend to be less sensitive to changes in the stimulus set than had been expected. The expectation of change in perceptual structures was based on: 1) the prior research finding that the underlying structure of the perceptual space for a set of products can be related to the characteristics of the stimuli, and 2) the implication that perceptual structure should be altered if the set of attributes represented by the stimuli changes.

The experimental treatments consisted of three different sets of brands: 1) a "standard" set of seven non-diet brands and one diet brand, 2) the previous set of stimuli with one non-diet brand replaced with another diet brand, and 3) a set consisting of four of the "standard" set non-diet brands, the "standard" set diet brand, and diet analogs of three of the non-diet brands. Similarity-dissimilarity ratings for all pairs of brands were based on sensory evaluations and on use of brand names.

The following questions and comments on this research appear to be relevant.

1) The justification presented for the measure of change in perceptual structure used in this study is that if the perceptual structure is stable, spatial relationships between stimuli should remain unchanged. Operationally, however, the authors rely on the reflection of this logic, i.e., if spatial relationships are unchanged, then perceptual structure is stable. The strength of this latter statement is questionable. If, for example, a new, independent and important, dimension were added to the perceptual space through inclusion of a new stimulus, the former brands may be perceived uniformly on this new dimension and their spatial relationship with each other remain unchanged. The authors' investigation of changes in dimensional salience may address this issue, but the consideration of two-dimensional solutions in each case could tend to suppress the observed effects of additional attributes being considered.

2) It was expected that the substitution of three diet brands for three non-diet brands would induce a change in the space to include a diet/non-diet dimension. It should be noted, however, that the diet/non-diet dimension was already represented in the "standard" set of stimuli. Therefore, the experimental treatments may not have added a new dimension but, instead, altered the distribution of stimuli on an existing dimension. In

light of the stated reasons for expecting change, i.e., alteration of the set of attributes represented, one might expect no change in perceptual structure.

3) Changes in the stimulus set did not result in differences in spatial relationships or dimensional saliences of the configurations based on brand evaluations. The corresponding configurations based on sensory evaluations did exhibit difference, however ( $r = .41$ ). The stability of configurations based on brand evaluations could be explained by the comments under (2), above. That is, the inclusion of Tab (a recognized diet brand) in the standard set had already introduced the diet/non-diet dimension, and substitution of more diet products did not affect the dimensionality. For the results with sensory evaluations, one could conjecture that the diet dimension was not detected in the "standard" set because only one of the eight brands was a diet drink, but increasing the number of diet drinks to four enabled the respondents to identify this dimension. If this were the case, the multiple substitution did add a previously undetected perceptual dimension, and the perceptual configuration changed significantly, as expected.

4) The reported results do not appear to support the general conclusion that perceptual structure is unchanged by alterations in the stimulus set. While configurations based on brand evaluations remained stable, those based on sensory evaluations did not (Table 2).

The authors have identified a question which should be of great interest to researchers scaling perception on the basis of paired comparisons and is worthy of further attention. The concern expressed in this review about the strength of the experimental treatment could be addressed through an additional analysis with Tab deleted from the "standard" stimulus set. The divergence in results based on brand and sensory evaluations also indicates a need for greater understanding of the effects of data collection method under various conditions. The authors' inclusion of both forms of presentation in this research is praiseworthy.

"Covariance Bias of Thurstone Case V Scaline as Applied to Consumer Preferences and Purchase Intentions"

This research [1] documents the existence of a possible problem in the use of the Thurstone Case V method of generating a unidimensional preference scale for a group of respondents. The covariance addressed in this research is the relationship between preferences expressed for two objects. The necessary assumption that these covariances are equal for all pairs of objects may be violated when data from a group of respondents are scaled. The primary reason for the violation is the possible homogeneity of perception combined with heterogeneity of preference. If objects A and B are seen as very similar by all respondents, there is likely to be high covariance in their preference ratings since a person's preference rating for one of them is likely to be similar to the rating for the other. Objects which are seen as dissimilar may have negatively correlated or uncorrelated preference ratings. Therefore, one could not expect the covariance for all pairs of objects to be equal.

In the "Conceptual Analysis" section, the authors

appear to argue that covariance is related to the proportion of times one object is chosen over another through an example in which one object would dominate a similar one because their momentary values would tend to move together. While this example may apply to scaling of data collected on several occasions from one respondent, it is not clear that it applies to analysis of data from several respondents on one occasion each. Two similar objects may have high covariance across a set of respondents, but it is not clear why this similarity and covariance should lead to the choice probability for this pair being any different than that for two dissimilar objects. In other words, if the preference levels for two similar objects are highly correlated across respondents, why should this lead one to expect that either would be preferred over the other with probability greater than .5? This reasoning is an important part of the paper, since it led to the expected nature of the effects of similarity on scale results.

The relationship between difference in transformed observed and predicted proportions and a measure of similarity was investigated. The following comments are related to this set of analyses.

1) The conceptualization of the problem was based on heterogeneity of preferences. It is a strength of this work that the expected effects were investigated on two data sets with different patterns of preference heterogeneity. While heterogeneity (as measured by average subject correlation) was not linearly related to the correlation of bias ( $z_{jk} - \hat{z}_{jk}$ ) and similarity, the direction of results supports the authors' contentions.

2) For each data set, subjects were divided into preference segments on the basis of "rank order correlations between individual subjects and a group scale." Additional information about this procedure would have been useful. It is not clear how the group scales can be developed without first identifying the individuals in the groups.

3) The measure of dissimilarity was based upon the physical characteristics of the objects (examples given as weave, pattern, lining, and variety of color). Each characteristic was coded as a dummy variable indicating its presence or lack thereof in each object. This set of dummy variable values appears to be a rather weak representation of the characteristics of an object, however. There may have been considerable variation in the colors and number of colors offered, for example. The objects by attributes matrix of dummy variable values was then submitted to principal components analysis. For the drapery data, the four attributes were reduced to two components explaining 79.4 percent of the variance. The concept of variance may take on a different meaning with 0-1 data, and the proper interpretation of the percent of variance explained by the components retained is not obvious. The Euclidian distance between designs on the components was used as the measure of similarity. For the above reasons, the measure appears to be "noisy." It is difficult to argue that the similarity biased results in a positive direction, and to the extent that the measure was noisy, the results were artificially weakened. It is possible, therefore, that a more precise similarity measure would have led to even stronger results.

4) The authors' point that subjective dissimilarity data would be preferable is a good one. The results may have been even stronger with such a measure.

5) The discussion of implications for relative versus "absolute" affective values is somewhat confusing. The authors seem to be saying that covariance bias will

result in inaccuracy in the precise location of points on the scale but that rankorders are not likely to be affected.

The importance of this paper is related to the extent to which Thurstone Case V scaling is actually used. The authors have documented the existence of a problem and suggested means of reducing its impact. It would have been useful to have the size of the effect of covariance bias shown in a table comparing scale values for all subjects and for the segments. Such a presentation would have demonstrated the importance of the problem, as well as its existence.

The authors argue that covariance bias applies also to models using the logistic function or the arcsin transform distribution. Many recent research efforts have made use of these functions (the logistic is frequently used on conjoint analysis, for example), and future efforts to determine the effects of covariance bias in these models should be of substantial interest.

#### "Market Share Estimates Based on Conjoint Analysis of Concepts"

This paper [3] proposes a logistic model for estimating market shares for alternative concepts whose overall utilities have been specified through conjoint analysis. The model is deterministic at the individual level, in that it is assumed that each person will choose the concept with highest overall utility to that person. The problem is viewed as stochastic at the aggregate level because of the heterogeneity of utilities across individuals. The model computes the probability that each concept is preferred to all others through use of the logistic distribution and translates these probabilities into market share estimates.

The following comments are related to the model and discussion presented in the paper.

1) In their introductory section, the authors discuss two types of factors that influence consumer preferences. The first consists of attributes and attribute level utilities as modeled in conjoint analysis. The second type of factor is external to the product itself and consists of use context, needs satisfied, institutional constraints, etc. It is commendable that the authors have considered factors in the choice process which are external to the constructs normally considered in models of preference. The perspective offered, however, seems to consider the product-related and "environmental" factors as affecting preference simultaneously. It is possible to view the determination of preference as a two-level process in which preference is determined by attribute utilities which, in turn, depend upon the environmental factors. The two views differ in that in one case the environmental factors are considered as additional explanatory variables for preference, and in the other they are considered as explanatory variables for variation in attribute utilities. The implication of the latter view for measurement in marketing research is that it suggests taking attitudinal measures within environmental scenarios. The alternative is to take global attitudinal measures (i.e., across environmental settings) and add environmental factors to these in order to explain preference.

2) The authors point out that one of the assumptions necessitated by use of the logistic model is that the correlation between difference scores (i.e.,  $(U_i - U_j)$  and  $(U_i - U_k)$ , across individuals) is equal to one-half for all combinations of pairs. It is then shown that this assumption holds when the Thurstone Case V assumptions hold ( $\sigma_i^2 = \sigma_j^2$  and  $\rho_{ij} = \rho_{ik}$  for all  $i, j$ ,

and k). The previous paper addressed these Thurstone Case V assumptions and concluded that they may not hold for preference data from a group of individuals. The research reported by Huber and Sewall [1] would therefore open to question the use of the logistic model suggested in this piece.

3) The model is described as providing a practical vehicle for estimating market shares that would be captured. It is based on an individuals by concepts matrix of overall utilities, and the fundamental assumption is that each individual will choose that alternative for which his utility is highest. It appears, however, that the use of an aggregate stochastic model is not necessary, and to the extent that the model does not fit the data, predictive accuracy is lost. A simpler approach would be to simply conduct a "simulation" in which choice is predicted individually for each person on the basis of his utilities. This procedure requires no more data since it is just a comparison of the entries in each row of the utility matrix. Market share can then be predicted as the proportion of individuals for whom each concept has the highest utility. One of the advantages of the simulation approach is that market segmentation implications can be explored directly through consideration of other descriptors of the individuals. This can be done either by dividing respondents into groups on an a priori basis and estimating market shares for the individual groups or by comparing the respondents who prefer the different concepts on other explanatory variables. The use of an individual simulation also allows the researcher to weight respondents by usage level.

One implication derived from the model is that a concept with a lower mean utility and larger variance may capture a larger market share than a concept with higher mean utility and smaller variance. This finding precludes use of aggregate mean utilities of concepts to predict market share, since correlations in utilities are not considered. It does not, however, appear to preclude the individual by individual type of prediction suggested above.

One of the strengths of this paper is the comparison of the logistic model with other models. This type of discussion is very useful to a field in which many research efforts appear on the surface to be unrelated.

#### Conclusions

These three papers reflect positively on this conference in that they investigate issues of pragmatic importance in marketing research. Although they may not be elaborate or comprehensive enough in their present states to be published in major journals, they do represent a contribution by helping to focus interest on specific problems of which current users of these methods may not be aware. They demonstrate one of the most important benefits of this conference in that developing streams of research and focused investigations of specific issues may be exposed and discussed among individuals with common interests. This process may be much more beneficial to the development of research methods in consumer behavior than the alternatives of little exposure or publication in specialized journals read by only a fraction of those potentially interested.

The problem area of developing and refining methods of measuring and modeling attitudinal constructs and using the results to forecast market behavior is filled with specific issues. Therefore, it is difficult to suggest a general direction for future research. If one takes a bit broader perspective, however, a rather fundamental question seems appropriate. Most perception and preference modelling efforts are directed at what

could be termed "global" constructs, e.g., preference for brand X. Are researchers missing a possible opportunity for greater understanding and predictive ability by paying insufficient attention to the purchase/usage context? One could hypothesize that the usage context will play an important role in affecting the determinants of behavior for many products, including those studied in these papers. A person's preference for sheets, for example, may be affected by whether they are to be used in the master bedroom, children's room, or a guest bedroom. At present, however, it appears that little is known about methods of identifying purchase/usage contexts of general importance and incorporating these contexts into measurement methods. A better understanding of the importance of context and means to operationalize it could suggest modification in method, lead to increased explanatory and predictive power (and, therefore, understanding), and hold implications for creative brand management strategy. Although these comments are not intended to deprecate the types of research efforts reported in this session, they do represent a call for a broader perspective and increased attention to a set of potentially important explanatory factors.

#### References

Joel Huber and Murphy A. Sewall, "Covariance Bias of Thurstone Case V Scaling as Applied to Consumer References and Purchase Intentions," Advances in Consumer Research, Vol. VI.

James McCullough, Charlene S. Martinsen, and Linda Scurman, "Changes in Consumer Perceptions: The Impact of Testing Conditions on Perceptions of Branded Products," Advances in Consumer Research, Vol. VI.

James B. Wiley and Robert Bushnell, "Market Shares Estimates Based on Conjoint Analysis of Concepts," Advances in Consumer Research, Vol. VI.

AN EMPIRICAL EXAMINATION OF ALTERNATIVE MODELS  
FOR PREDICTING CONSUMER UTILIZATION OF TWO  
CREDIT CARD SYSTEMS

Elizabeth C. Hirschman, New York University  
Rajendra K. Srivastava, University of Texas at Austin  
Mark I. Alpert, University of Texas at Austin

Abstract

Credit-facilitated purchase behavior is an increasingly important phenomenon in society. Within this context little empirical research has investigated predictive models of credit system preferences and utilization. This paper reports three models developed in a hierarchical fashion from demographic, mobility and benefit variables applied to users of two alternative credit card systems.

Introduction

Credit cards appear to serve three general functions in society. First, they may effectively ease purchasing on credit by reducing transaction time. Second, they may permit more immediate gratification of consumer needs and wants. Third, they have been said to serve as an economic stimulus and thus raise living standards. Credit cards have also been characterized as having some negative social effects. Among these are first, the stimulation of materialistic values and hedonism; and second, the creation of indebtedness among consumers which may lead to anxiety and bankruptcy.

Despite the normative debate concerning their relative utility, credit cards are an innovation with a long-standing history in the United States. Here credit cards have been introduced consecutively in three major forms: retail store-issued credit cards, "travel and entertainment" (Diners Club, Carte Blanche, etc.), "bank" cards (VISA, Master Charge). Each possesses different combinations of common and unique attributes. It would appear that by virtue of coming last the bank card is the most advanced of the credit cards. However, at the present time it shows no signs of eliminating its predecessors. That different consumer need-fulfillments are provided by each of the three forms of credit card is evidenced by the fact that many people carry two and sometimes all three of the different forms of cards, as well as multiple types within a particular form [Shay and Dunkelberg, 1975].

The credit card in all three forms may be usefully conceptualized as a service system for facilitating value exchanges. That is, the card, itself, is simply a symbol which permits the carrier to effect monetary transactions. The three basic credit card systems available at the present time, however, are all constrained with respect to the types and value of the transactions which may be conducted. For example, a credit card issued by a department store may generally only be used to conduct transactions at units of that store. A credit card issued by a bank, while usually enjoying a wider geographic and inter-retailer acceptance than a typical retail store credit card, is still not accepted by all retail outlets, particularly those retail chains having their own nationwide credit systems [i.e. Penneys, Sears]. Further, both of these credit systems are limited in their utility to the user by issuer-imposed "ceilings" or "credit limits" on the absolute monetary value of the exchanges that may be made with them.

Despite the similarity of their constraints, the retail store credit card system and the bank credit card

system would appear to provide the user with a dissimilar set of benefits. For example, the fact that a retail store issues its own credit card may provide the user with a perceived enhanced ability to negotiate with the store in exchanging or returning merchandise and in rectifying incorrect billing, [Shay and Dunkelberg, 1975].

Conversely, the bank issued credit card system places an intermediary [i.e. the bank] between the purchaser and the retailer, thereby reducing the direct negotiating power of the buyer. However, as mentioned earlier, this potential disadvantage of the bank card may be offset for some users by the fact that the bank card is accepted as a valid means of exchange at a much wider variety of retail outlets, as well as having almost global recognition and acceptance.

It would appear both theoretically and pragmatically useful to determine if this hypothesized benefit structure is so perceived by the consumer. Of perhaps even greater utility would be the determination of whether or not an individual's credit-card-facilitated exchange transactions were influenced by the way in which the benefits attributed to a credit card were evaluated. In other words, the theoretical research question becomes one of ascertaining whether persons who place differential importance on certain credit card attributes also favor and utilize different types of credit card systems for facilitating their exchange processes.

A high correlation between perceptions of credit card system features evaluations and credit card choice, of course, paves the way for a benefit segmentation approach to both the understanding and promoting of a given credit system. However, in the interest of parsimony, it would be desirable if such a benefit-structure segmentation proved superior to segmentation derived from a simpler and more easily applied basis such as demographics. Demographic variables have already been demonstrated effective in describing and predicting users of bank-issued credit cards [Adcock, Hirschman, Goldstucker 1976; Goldstucker, Hirschman 1977; Hirschman, Blumenfeld, Tabor, 1977; Matthews and Slocum, 1969; Plummer 1971; Slocum and Matthews 1975; Wiley and Richard, 1975], as well as retail store-issued credit cards [Russell, 1975; Mateer, 1969; Shay and Dunkelberg, 1975]. Therefore, to be considered most useful in a pragmatic sense, benefit segmentation should be shown to serve either as a replacement for or complement to the traditional segmentation base of demographics.

To enhance the potential utility of such a segmentation approach, one should investigate reported purchase behavior, [rather than reported preferences] and a situation in which the individual is confronted with a genuine choice between alternative credit card systems. One shortcoming of much of the previous research has been that the descriptive profiles developed have been based upon an individual's "usual" mode of payment for credit purchases or, even more often, simply upon possession of a certain type of card. Few attempts have been made to measure the correlation between possession



and usage. Second, in many instances no control has been made for different situational contexts. For example, an individual may "usually" use a bank-issued credit card, but in department store purchase situations may prefer (or be forced) to use a retail store issued card. These possible limitations in some of the prior research on credit cards may cast some doubt on the validity of the findings reported.

Thus, it would appear that to address adequately the problem of measuring the validity of a demographic-based credit card system choice model versus one based solely or incrementally on perceived benefits two conditions must be present--the individual's purchased behavior must be studied while controlling for the usage-context, and the individual must be confronted with a genuine choice as to credit card payment mode.

The purpose of this research is therefore three-fold. First, an investigation will be made to determine if persons who primarily use bank-issued credit cards when they have a choice of these versus retail store-issued cards can be effectively discriminated from those who prefer to use the latter, using standard demographic variables. Second, we shall investigate whether more effective discrimination and prediction can be obtained by augmenting basic demographics with particular demographics deemed relevant to choices of credit systems, such as mobility. Third, we shall investigate what, if any, additional explanatory and discriminatory power may be gained by utilizing information concerning the relative importance of various credit system benefits to those individuals choosing alternative systems. The addition of non-demographic variables should be advocated only to the extent that they provide significant explanatory and predictive strength to the process of understanding consumer choices and their implications for (in this case) marketing and credit management.

#### Methodology

Data Gathering: Data for this research were drawn from a survey of department store customers who were intercepted on a random basis by professional interviewers and administered a structured questionnaire. The questionnaire dealt with purchases made at the store, mode of payment used, attitudes concerning various credit instruments, store patronage, and a detailed set of demographic characteristics. The survey was conducted in two major southwestern cities from March 21 to April 3, 1977. One of the survey cities has a population of slightly under 2,000,000; the other has approximately 500,000 inhabitants.

In the first and larger city 3,000 interviews were conducted, while 1,225 were gathered in the second city. A department store chain operating stores in both cities cooperated in the research by allowing interviews to be conducted with exiting customers. All interviews were conducted on a random intercept basis at the major store entrances during all hours of normal store operations. Interviews were conducted at five branch stores in the first city and one branch store in the second city.

Because certain quotas of purchasers/non-purchasers and bank card users/non-users had to be met and respondent screening used to meet these quotas, the demographic data are not representative of this department store chain. Given the focus on credit system choices, cash-users are underrepresented here. This qualification, however, does not inhibit the utility of the data for the purposes intended here.

#### Questionnaire

The questionnaire used to generate the data for this research consisted of an extensive set of structured

questions designed to investigate several aspects of shopping and credit-related activities. One of the portions of the questionnaire relevant to this research consisted of two questions designed to qualify the respondent as a person who usually used either the bank card or the retailer-issued credit card for making purchases at the store, and who had prior experience in making credit purchases at the store. These questions are given in Exhibit 1.

#### EXHIBIT ONE

#### VARIABLE MEASUREMENT

##### Criterion:

1. How do you usually pay for your purchases at this store?
 

<input type="checkbox"/> a. Cash	<input type="checkbox"/> c. Bank Card
<input type="checkbox"/> b. Store Charge Card	<input type="checkbox"/> d. American Express Card
  
2. Have you ever charged a purchase at this store by using a credit card other than the store credit card?
 

<input type="checkbox"/> a. Yes	<input type="checkbox"/> b. No
---------------------------------	--------------------------------

##### Demographic Predictors

Respondent's Age  
 Respondent's Education  
 Spouse's Education  
 Household Annual Income  
 Sex  
 Race  
 Life Cycle Stage:  
 Young, Married, No Children  
 Married, Children Under 6  
 Married, Children Over 6  
 Older, Married, No Children At Home  
 Older, Single

##### Mobility Predictors

Residency Period:  
 Under 1 Year  
 One to Three Years  
 Three to Six Years  
 Six to Ten Years  
 Over Ten Years  
 Resident/Non-Resident

##### Benefit Predictors

1. Are the following characteristics of very much importance, of moderate importance or of very little importance to you in using a charge card?
  - a. Ability to use at a wide variety of stores.
  - b. Ability to use all over the country.
  - c. Ability to easily return merchandise.
  - d. Ability to use as a means of identification.
  - e. Ability to obtain easily.
  - f. Ability to utilize additional credit plans.
  - g. Reputation or prestige of the card.
  - h. Interest Rates.
  - i. Ability to straighten-out incorrect billing.
  - j. Ability to replace if stolen or lost.
  - k. Ability to have one or two larger bills rather than a lot of small bills.

A second portion of the questionnaire relevant to this research was concerned with the respondent's demographic characteristics. These demographics were divided during the model-building part of the study into two groups. The first group consisted of the "basic" demographic traits of age, education, income, sex, race and life cycle stage, and are shown under the heading "Demographic Predictors" in Exhibit One. A second group of demographic variables--length of residency in

the market and resident/non-resident status were held out of the basic demographic model, as it was anticipated that they would possess significant incremental predictive ability, even though they are infrequently used in demographic segmentation studies. These are shown under "Mobility Predictors" in Exhibit one.

Third, a set of eleven attributes relevant to the selection of alternative credit card systems was included in the questionnaire. The respondent was asked to tell if each attribute was of very much importance, moderate importance or little importance in using a credit card. Responses were scored in reverse order, 1, 2 or 3, respectively. These are shown in Exhibit One, also, under the heading "Benefit Predictors."

This set of benefits was arrived at by conducting a series of ten focus group sessions in the two markets prior to the in-store interviewing phase of the research. The topic at each of the sessions was the use of credit and credit cards. The benefits [i.e. relevant attributes] identified by the various focus group sessions were largely consistent from group to group regardless of group composition. Thus, the benefit list developed is believed to possess some degree of generalizability.

#### Hypotheses

Three hypotheses were generated at the outset of the research. First, it was anticipated that since there are few, if any, differences in the demographic qualifications imposed by issuers of bank cards and retail store credit cards on persons applying for these cards, it follows that there should be few demographic differences between the groups using these cards. In fact, since persons belonging to either the bank or store card user groups possessed both types of cards, demographic variables should be of little value in differentiating between them. Further, since the retail store surveyed appeals largely to higher income customers, little covariation between card choice in this context and demographic characteristics would be expected.

Second, it was anticipated that a shorter residency period [implying greater mobility] and visitor status should be useful discriminators between bank card users and store-issued card users. The reasoning behind this hypothesis was that persons who were from out of town or recently arrived would prefer a credit card that could be used across a wider geographic area, and also have less store loyalty than longer-term residents; both these factors should contribute to a heightened likelihood that such persons would use a bank card as their usual purchase mode.

Third, it was hypothesized for the reasons set forth in the introduction to this paper that the benefits of ability to return merchandise and straighten out incorrect billing would be more highly valued by persons using store-issued credit cards; while the benefit of ability to use at a wide variety of stores, ability to use all over the country and ability to consolidate bills would be more highly valued by persons using bank credit cards.

These three research hypotheses are stated below:

H1: There will be no demographic differences between persons who prefer to use bank cards, ["bank card users"], and those who prefer to use store-issued credit cards, ["store card users"].

H2: There will be discriminating mobility differences between bank card users and store card users.

H3: There will be discriminating benefit importance differences between bank card users and store card users.

For statistical analysis, hypotheses were restated in the 'null' form.

#### Analysis and Findings

To test the predictive and explanatory validity of the demographic, demographic plus mobility and demographic-mobility-benefit models, a group of respondents who possessed a bank card and a card issued by the department store whose customers were surveyed was selected. Since the store's acceptance of bank cards was a recent event, we narrowed the analysis to include only those persons who had used both types of credit cards to charge purchases at this store chain. This procedure would assure first, that persons who chose to use the retail card did not do so simply because of ignorance of the bank card's acceptability at the store rather than a preference for retail credit cards, and second, that the person was not constrained by lack of possession of one of the two relevant credit cards. Bank card users were defined as those who had used both types of cards but indicated that they usually used bank cards to pay for their purchase at the store. Retail card users were similarly defined as persons who had used both types but usually paid for their purchases with their retail card.

Having obtained a group of respondents who were qualified to choose between using their store-issued card or their bank card, we could be sure that the choice involved a conscious and unconstrained decision. A series of multiple discriminant analyses were run to test the discriminating ability of hierarchical combinations of demographics [Model I], mobility [Model II], and benefits sought [Model III]. By adding mobility to demographics, we could see whether additional explanatory power could be obtained by considering the impact of recent movement to the city, with other demographic variables "held constant" by the discriminant function. By adding benefits sought to this enriched model, we could see whether persons with similar demographic and mobility profiles might have differential probabilities of choosing to usually use a store-issued versus a bank card, when considering benefits sought from credit systems. To the extent that benefits are correlated with demographics and/or mobility, the incremental contribution of these variables would be lessened, once the predictive ability of the enriched demographic/mobility model had been obtained. However, if additional predictive validity could be obtained by considering attitudes not already "explained" by the earlier variables, some additional conceptual and pragmatic strength could be claimed.

Table 1 presents a summary of the results of testing the three models in the form of linear discriminant analysis, using the BMDP7M stepwise discriminant analysis with the "jackknife" method of cross-validation. There were 242 respondents who had used both types of credit systems at the intercept stores, and had supplied complete answers to the questions measuring the variables that were to be used to attempt to classify them. Persons who had not used both types of cards, or who had missing data were eliminated to prevent making specious comparisons. Of these 242 respondents, 119 indicated that they normally used a bank card at the intercept store, and 123 indicated a loyalty to the retail card of that store. While these figures provide approximately equal group sizes for the discriminant analysis, it is important to note that the percentage actually using bank cards in these stores would be substantially lower than 50%, since this sample was trimmed to include those who had actually used both types of cards.

TABLE 1

## COMPARISON OF MODELS - DISCRIMINANT ANALYSIS

	Wilks' lambda	"R <sup>2</sup> "	Rao's V	$\Delta$ Rao's V	$\Delta$ °F	% Correctly Classified ("Jackknifed")	% Correctly Classified <sup>1</sup>
Model 1: Demographics	.9175	.0825	21.079	21.079 <sup>a</sup>	10	56.6	56.6 <sup>a</sup>
Model 2: Demographics plus mobility	.7486	.2514	80.864	59.725 <sup>b</sup>	2	68.6	12.0 <sup>b</sup>
Model 3: Demographics plus mobility plus credit card benefits sought	.5578	.4422	198.872	118.068 <sup>b</sup>	11	78.1	9.5 <sup>b</sup>

P<sub>a</sub> < .05P<sub>b</sub> < .01

N = 242 in both groups combined

Table 1 indicates that the demographic variables alone were able to discriminate between "bankcard-loyal" users, as the Rao's V statistic for Model 1 was significant at beyond the .05 level. While statistical significance for demographics was found, the practical significance was marginal, as indicated by two factors. Using a regression approach to discriminant analysis [Green & Tull, 1975], one can view the discriminant problem as regression with a binary dependent variable. For this 2-group problem, it can be shown that "R<sup>2</sup>" = 1 - Wilks' Lambda. Accordingly, one measure of the marginal practical significance of the discriminant function utilizing demographics is its Wilks' Lambda of .9175 or "R<sup>2</sup>" = .0825.

A more traditional interpretation of practical significance is the percentage of group members who could be correctly classified by the predictor variables. To avoid the upward bias that would occur by using the same data to generate and then validate the discriminant function coefficients, it is necessary to cross-validate the function(s) by holding out data from the parameter-

generation phase and then classifying these cases using parameters from the estimation sample. Rather than "waste" considerable data by running a conventional split-sample procedure, we used the "jackknife" option of the BMD Program. Under this option, one case is held-out while parameters are estimated from the N-1 data points. These parameters are used to classify this case, and a "hit" or a "miss" is scored. The case is then pooled with the others, while another one is extracted for validation and parameters are again estimated with a "new" N-1 cases. Repeating the procedure N times provides a complete test of the classification ability of the variables without ever using any to calculate their "own" function [Crask and Perreault, 1977]. In addition the weights for each estimate can be averaged to provide the same kind of stability that would have been sought by using the entire sample without "throwing away" data for cross-validation.

Our interest at this point is in describing the predictive validity of the demographic model, while the discriminant weights will later be used to examine explanatory power using Table 2, which describes the discriminating variables. The demographic variables alone correctly classified 56.6% of the users in the "jackknife" validations, while 50.0% would have been expected according to the proportional-chance criterion [Morrison, 1969]. While this represents a statistically significant percentage correctly classified, the large proportion of misclassification errors parallels the low R<sup>2</sup> estimate and indicates that demographics alone would not provide very effective classification variables for credit card system choices [Hypothesis 1].

On the other hand, adding two quasi-demographic variables that measure consumer mobility significantly improved the discriminatory power of the set of demographic variables [Hypothesis 2]. Model 2 improved the "R<sup>2</sup>" to .2514, and the change in Rao's V was significant at beyond the .01 level. Accordingly, the correct classification percentage improved 12.0% to 68.6% a change which was also significant at the .01 level. Explanatory power of the relationships was also improved, as will be noted below.

Model 3 augmented the two previous models by incorporating the stated importance which respondents attached to benefits provided by alternative credit card systems. Adding these benefits sought from credit cards

<sup>1</sup>Unlike the "change in Rao's V" test for the change in the statistical discriminatory abilities of models which are formed by adding variables to existing models, there exists no straightforward test for the significance of incremental practical significance, or percentages correctly classified by two models. However, it is possible to construct such a test by recognizing that the cross-validated predictions of various models formed incrementally may be treated as "repeated measures" or correlated observations in an Analysis-of-variance design. The dependent variable becomes dichotomous ("0" for misclassified, "1" for correctly classified), but simulations have shown that ANOVA is appropriate for dichotomous as well as continuous measures, provided the samples are reasonably large (upwards of 30 to 50 subjects per group, depending on the mean percentages scored as "1"). Table 1 thus summarizes the ANOVA results for these comparisons of changes in "success-rates" for pairs of correlated observations of the success/failures of the four models.

TABLE 2

DESCRIPTION OF MODELS - DISCRIMINANT ANALYSIS						
	T-Ratio for Discrim. Wt. ÷ Std. Error					
	Bank Card Mean	Retail Card Mean	2-Sample T-Ratio	Model 1	Model 2	Model 3
<u>Model 1: Demographics</u>						
	n=119	n=123				
Respondent's Age	3.62	3.74	-.899	.982	1.334	.664
Respondent's Education	5.53	5.35	1.337	.422	-.200	-.853
Spouse's Education	5.85	5.59	1.987 <sup>a</sup>	1.569 <sup>a</sup>	.920	1.275
Household Income	5.24	5.50	-2.001 <sup>a</sup>	2.369 <sup>a</sup>	-2.143 <sup>a</sup>	-2.072 <sup>a</sup>
Sex (M=1, F=2)	1.22	1.14	1.636	1.721	.983	.200
Race (white=1, NW=2)	1.03	1.08	-1.590	-1.770	-1.517	-.908
Married, under 35, no children	.26	.19	1.373	.126	-.084	.427
Married, children under age 6	.25	.22	.595	.200	-.032	.435
Married, children over age 6	.43	.44	-.295	.000	-.187	.389
Married, over 35, no children at home	.06	.15	-2.250 <sup>a</sup>	-.597	-.632	.110
<u>Model 3: Demographics &amp; Mobility</u>						
Length of Residence in Area	2.23	3.99	-8.112 <sup>b</sup>		-3.711 <sup>b</sup>	-3.415 <sup>b</sup>
Visitor (0=No, 1=Yes)	.31	.03	6.19 <sup>b</sup>		1.167	1.157
<u>Model 3: Demographics &amp; Mobility &amp; Benefits Sought*</u> (1=Much Importance, 3=Little Importance)						
Use of Variety of Stores	1.16	1.28	-1.959 <sup>a</sup>			-.603
Use all over country	1.28	1.46	-2.134 <sup>a</sup>			.381
Ease of Merchandise Return	1.96	1.40	5.500 <sup>b</sup>			2.977 <sup>b</sup>
Use of Identification	1.81	1.68	1.182			.919
Ease of Obtaining	2.26	2.07	1.815			-.484
Utilize additional Credit Plan	2.56	2.22	3.445 <sup>b</sup>			2.394 <sup>a</sup>
Reputation/Prestige of card	2.29	2.30	-.138			-2.065 <sup>a</sup>
Interest Rates	2.22	2.04	1.566			-1.645
Bill Correction	1.79	1.50	2.702 <sup>b</sup>			1.925
Ability to replace if lost/stolen	1.50	1.44	.591			-1.740
Consolidate Billing (0=No, 1=Yes)	.42	.11	5.736 <sup>b</sup>			6.202 <sup>b</sup>

\* Benefits scored in reverse order. The lower the mean, the greater the importance

$p_a \leq .05$  (2-tailed)       $p_b \leq .01$  (2-tailed)

Note "-s" mean the higher the variable score, the more likely the person is a storecard user.

substantially improved the ability to discriminate retail from bank card users, [Hypothesis 3]. The change in Rao's V was significant beyond the .01 level, and the "R<sup>2</sup>" improved to .4422. Further, the percentage correctly classified in the jack-knife cross-validation jumped another 9.5% [significant beyond the .01 level], giving the combined model the ability to classify correctly 78.1% of the sample. The explanatory power of the three models will be discussed next.

Table 2 provides data on the abilities of particular variables within each model to discriminate group membership, in a univariate sense, as well as their individual contributions to the overall Model 3 in a multivariate sense. Because of the presence of intercor-

relations among some of the predictor variables, one must use caution in interpreting the relative importance of variables based on the traditional criterion of significance of the discriminant weights.

However, combining this information with the univariate 2-group comparisons between mean scores for variables (column 3) may provide some explanation for the credit card choices and marketing implications of this behavior.

For example, according to the ratio of discriminant weight divided by its standard error, the only demographic variable that provides significant discriminatory power, holding constant other demographics,

mobility, and benefits sought, is household income. All things being equal, the higher the income, the more likely a respondent was to choose the retail card. Other data in the survey also show that retail card users were more frequent shoppers at this "top-line" department store than were bank card loyal users, which correlates with the importance of income as a discriminator. In fact, when frequency of shopping at the store is controlled for (by adding it to the discriminant function), the effect of income is no longer statistically significant.<sup>2</sup>

In a managerial sense it is probably worth knowing that those preferring to use the retail card have higher income levels and are more frequent shoppers at the store. Heavy users of the retail credit cards "make good customers," especially when it is known that their income is high enough to lead to low credit losses. At the conceptual or theoretical level, however, the fact that bank card users who are equally frequent customers of the store have similar incomes to the retail card users suggests that there is nothing inherently discriminating about retail vs. bank cards' appeals to various income classes, when purchase rates are "controlled." Nevertheless, multicollinearity inhibits making these kinds of theoretical statements from these data, as the stability of the coefficients and the ability to "control" other variables is open to question.

For example, the significantly negative discriminant weight for the importance of reputation of prestige of the card indicates that, *ceteris paribus*, the lower the score on that variable, the more likely one is a bank card user. Given the coding of importance, a low score implies high importance for the prestige variable among bank card users. Whether this is in compensation for their lower income or, holding income constant, a greater need for the prestige of bank cards, or an artifact of unstable (due to multicollinearity) discriminant weights remains to be established in future research.

For the two mobility variables added to demographics to make Model 2, the univariate and multivariate inferences are more consistent. Retail card users as a group are longer-time residents in the survey city, as indicated both by the comparison of group means as well as the significantly negative discriminant weight (implying a lower score for the bank card users). A separate analysis showed that 31% of the bank card users were visiting the city (which also correlates with their lower frequency of store shopping) vs. 3% of the retail card loyal users. The discriminant coefficient is marginally significant in the same direction (1-tailed  $\alpha = .06$ )

As is shown in Table 2 several of the benefits sought from credit card systems do possess a significant relationship to which type of credit card was used. As put forward in Hypothesis Three it was anticipated that persons using bank cards would place more importance on the attributes of ability to use a card at a wide variety of stores, the ability to use a card all over the country, and the consolidation of billing. It was also expected that persons who used the store-issued card as their usual mode of payment would place more importance on the ability to easily return merchandise and to straighten out incorrect billing. These expectations were entirely confirmed by the univariate t-tests with significance levels mixed between  $p \leq .05$  and  $p \leq .01$ .

<sup>2</sup> For brevity, the analysis controlling for frequency of shopping is not reported here.

In addition to these expected differences, one which was not anticipated also appeared. This was the significantly greater emphasis users of store-issued cards placed upon the ability to utilize additional credit plans, [ $p \leq .01$ , 2-tailed test]. Given the presence of credit limits on bank cards, this finding is intuitively appealing.

When the benefits sought were added to the discriminant model containing the demographic and mobility-related variables, many of these univariate differences disappeared. This would appear to indicate that some benefits sought are closely associated with demographic or mobility-related characteristics of the individual, as we suggested in Hypothesis 2. Specifically, differences between bank card users and store-issued card users on evaluations of the ability to use a card at a wide variety of stores or all over the country were no longer significant once residency period and resident/non-resident status had been controlled for [which effectively "removed" mobility as a variable]. A significance level of between  $p = .10$  and  $.05$  was still found for the ability to straighten-out bills, and differences between evaluations of ability to easily return merchandise, utilization credit plans, and consolidated billing were still significant in the same directions at the  $p \leq .01$ , and  $.01$  levels, respectively.

#### Summary and Conclusions

This research was designed to investigate alternative models for predicting usage of retail store issued and bank issued credit cards. To develop these models a subsample was selected from the major sample on the basis of the following criteria: (1) that individuals possessed both types of cards, and (2) that individuals had prior experience in using both types of cards. The second restriction (made feasible by question 2, Exhibit 1) is especially relevant because a respondent may possess both types of cards, yet not be aware that bank credit cards could be used at the store under investigation. Of the three research hypotheses proposed:

- H1: Was rejected at the  $p = .05$  level. That is, there were statistically significant differences in demographic variables between users of bank cards and store-issued cards.
- H2: Was confirmed at the  $p = .01$  level. That is, there were discriminating residency differences between users of bank cards and store issued cards.
- H3: Was confirmed at the  $p = .01$  level. That is, there were discriminating benefit importance differences between users of bank cards and users of store-issued cards.

Examination of Tables 1 and 2 reveals that while the demographic variables are statistically significant, little practical significance (in terms of predictive power) can be gained from Model I. The addition of the two mobility variables in Model II provided both statistically and practically significant improvements in prediction. Similar improvements were also noted when variables reflecting the importance of benefits sought from credit systems, Model III, were added.

The individual variables that were significant in a multivariate sense (that is, by statistically controlling for the other variables) were: household income (demographic); length of residence in area (mobility); ease of merchandise return, utilization of additional credit plans, reputation/prestige of card, and consolidated billing (benefits sought). As discussed, the

direction of the differences in group means were in agreement with the expectations cited earlier in the paper.

The results provided by this study may possess both theoretical and practical utility, as they illustrate the pattern of benefits sought with respect to credit systems by two diverse types of credit card users in a retail purchase setting. This may pave the way for a benefit segmentation approach which will be generally applicable to credit-facilitated exchange processes.

On a more theoretical level, the findings from the study suggest a need to explore the linkage between credit card system benefits sought and demographic and life style related variables. Such research may serve to develop a more substantial theoretical base for understanding credit system choices and credit-facilitated consumer purchasing behavior.

#### References

William O. Adcock, Elizabeth C. Hirschman and Jac. L. Goldstucker, "Bank card Users: An Updated Profile," paper presented at the Association for Consumer Research Annual Conference, 1976.

R.V. Awh and D. Waters, "A Discriminant Analysis of Economic, Demographic, and Attitudinal Characteristics of Bank Charge-Card Holders: A Case Study," Journal of Finance, 29 (June, 1974), pp. 973-980.

Melvin R. Crask and William D. Perreault, Jr., "Validation of Discriminant Analysis in Marketing Research," Journal of Marketing Research, 14 (February, 1977), 60-8.

Jac L. Goldstucker and Elizabeth C. Hirschman, "Bank Credit Card Users: A New Market Segment for Regional Retailers," MSU Business Topics, 25, (Summer 1977), pp. 5-11.

Paul E. Green and Donald S. Tull, Research for Marketing Decisions, (Englewood Cliffs, N.J.: Prentice-Hall, Inc. 1975), pp. 453-56.

Elizabeth C. Hirschman, Warren S. Blumenfeld, and Dwight Tabor, "An Attempt to Use Respondent Sex as a Moderator for Profiling Bank Card Users," Proceedings Ninth Annual Conference, American Institute for Decision Sciences, Chicago, 1977, pp. 227-229.

William H. Mateer, The Checkless Society, MSU Business Studies, Bureau of Business and Economic Research, Michigan State University, East Lansing, Michigan, 1969.

H. Lee Matthews and John W. Slocum, Jr., "Social Class and Commercial Bank Card Usage," Journal of Marketing, 33 (January, 1969), pp. 71-78.

\_\_\_\_\_, "Rejoinder to 'Social Class or Income?'" Journal of Marketing, 36 (January, 1972), pp. 69-70.

Donald G. Morrison, "On the Interpretation of Discriminant Analysis," Journal of Marketing Research, 6 (May, 1969), 156-63.

Joseph T. Plummer, "Life Style Patterns and Commercial Bank Credit Card Usage," Journal of Marketing, 35 (April, 1971), pp. 35-41.

Thomas Russell, The Economics of Bank Credit Cards, New York: Praeger Publishers, 1975).

Robert P. Shay and William C. Dunkelberg, Retail Store Credit Card Use in New York, Studies in Consumer Credit, No. 4, Graduate School of Business, Columbia University Press, New York, 1975.

John W. Slocum, Jr., and H. Lee Matthews, "Social Class and Income as Indicators of Consumer Credit Behavior," Journal of Marketing, 35 (April, 1975), pp. 59-74.

James B. Wiley and Lawrence M. Richard, "Application of Discriminant Analysis in Formulating Promotional Strategy for Bank Credit Cards," in Advances in Consumer Research, Vol. II, Mary Jane Schlinger, Editor, Association for Consumer Research, 1975, pp. 535-544.

THE ROLE OF SEX ROLE SELF-CONCEPT IN MASCULINE  
AND FEMININE PRODUCT PERCEPTIONS

Linda L. Golden, University of Texas at Austin  
Neil Allison, University of Cincinnati  
Mona Clee (student), University of Texas at Austin

Abstract

The impact of respondents' sex role self-concept, sex, product use and self-esteem on masculine and feminine product perceptions is investigated for five representative sex-role stereotyped products. Sex role self-concept appears to be more important for feminine product perceptions than masculine product perceptions. However, the respondent's sex and product use is at least as important for product perceptions as sex role self-concept.

Introduction

Females have become the focus of intensive marketing investigations. However, attempts to differentiate working and non-working women, as well as women with traditional and non-traditional attitudes has not resulted in clear distinctions in behavior. An alternative approach would be to differentiate women on the basis of sex role self-concept. Further, consideration of sex role self-concept is not only germane for females, but relevant to the other 49 percent of our population: Males.

Sex role self-concept has been studied within the psychology literature in a wide variety of contexts ranging from success orientation to male and female attraction (Bem 1972, Bieliauskas and Milesell 1972, Deusch and Gilbert 1976, Feather and Rephelson 1973, Gordon and Hall 1973, Jordan-Viola, et. al. 1976, Kravetz 1976, Seyfried and Hendrick 1973, Shaffer and Wegley 1973, Tolor 1976). Marketing investigations, however, have been limited. Three studies were found in the marketing literature which operationalized some measure of sex-role self-concept. Gentry, Doering and O'Brien (1978) found weak support for a congruency between sex role self-concept and the use of products with a neutral sexual identity or a sexual identity opposite from their own. Fry, (1971) also found that behavior was consistent with masculine-feminine self image. And, Morris and Cundiff (1971) found that males with a relatively high feminine identity and a high level of anxiety expressed strongly unfavorable attitudes toward the use of hair spray (perceived as feminine).

Marketers have more extensively studied the relationship between consumer behavior and the global self-concept (Birdwell 1968, Dolich 1979, Grubb and Hupp 1968, Grubb and Grothwohl 1967, Hamm and Cundiff 1969, Hughes and Geurrero 1971, Landon 1974). The relationship of self theory to buying behavior is based on the assumption that there is an interaction of the buyer's personality with the symbolic image of the product purchased (Levy 1959). In the midst of varying methodologies and different focuses, some studies support the congruence while others do not. However, the potential relationship between self-concept and consumer behavior cannot at present be negated.

Focus upon a specific aspect of the self-concept for both the product and the respondent may reveal relationships undiscernable by global congruency attempts. For example, examination of the congruency between the respondent's sex role self-concept and the perception of a product as having masculine and feminine dimensions.

Following from self theory and buyer behavior it would be expected that there would be a congruency between sex role self-concept, product perceptions on masculine and feminine dimensions and product use. Further, the sex of the respondent may impact upon this congruency. A male who perceives himself as having primarily stereotypical masculine traits may tend to use products perceived as high in masculinity (the products may be used because they're perceived as masculine, or some may be perceived as masculine because they're used--no cause and effect relationship is implied). Similarly, a person whose self-perceptions contain large amounts of both masculine and feminine traits (androgynous) may use products perceived as being of either sex-type.

In instances where an individual's sex role self-concept does not agree with the sex (cross-sex typing) a congruency relationship is not so clearly predicted. Will product perceptions and use tend to be congruent with the sex of the respondent or the sex role self-concept? The self-esteem of the individual may provide insight into this question (again, no cause and effect relationships implied).

Self-esteem has been shown to be related to sex role self-concept (Spence, Helmreich and Stapp 1975). For both males and females, androgynous persons were highest in self-esteem, followed by high masculine low feminine persons, then low masculine high feminine persons, and lowest in self-esteem were persons with no strong identity with either sex. Although in this study the cross-sex typed persons were not the group with the lowest self-esteem, within cross-sex typed groups of people variances in self-esteem may help to explain the relationship between sex-typed product perceptions, product use and sex or sex role self-concept.

Purpose and Objectives

This is an exploratory study designed to investigate the influence of sex role self-concept upon product perceptions. In addition, this study investigates the influence of sex, product use and self-esteem upon masculine and feminine perceptions of products. Five products are included in the investigation: pocket knife, cuff links, nylon underwear, hair spray and key ring. These products represent masculine-typed products, feminine typed products and a neuter-typed (low masculine and feminine identity) product. Due to the exploratory nature of this research, the following research objectives are advanced in lieu of formal hypotheses:

1. To investigate the congruency between sex role self-concept and sex-typing of products;
2. To determine the individual and combined effects of sex, use and self-esteem upon sex-typing of products;
3. To determine differences in the relationships between sex role self-concept, sex, use and self-esteem for varying types of products, and
4. To investigate self-esteem differences across sex

role self-concept categories.

### Methodology

In order to determine the sex-typed perceptions of products to be tested, a pre-test sample of 80 male and female business students were administered a questionnaire listing 125 generic product categories ranging from nylon underwear to brief case. Under the name of each product were two scales: one for masculine product perceptions and one for feminine product perceptions. The horizontal five-point scales had extremes labeled extremely (masculine and feminine) and not at all (masculine and feminine). The purpose of this pre-test was to sort products into nine groups: 1) high masculine and high feminine, 2) high masculine and medium feminine, 3) high masculine and low feminine, 4) medium masculine and high feminine, 5) medium masculine and medium feminine, 6) medium masculine and low feminine, 7) low masculine and high feminine, 8) low masculine and medium feminine, and 9) low masculine and low feminine.

The results of this pre-test indicated that the products fell into one of three groups: high masculine-low feminine, low masculine-high feminine, medium masculine-medium feminine. In addition, respondents indicated difficulty in understanding that the mid-point on the scale was "moderate" (masculinity or femininity), not "neuter", and in being able to keep track of their frame of reference when both the masculine and feminine scales were presented together underneath the product. Due to these comments and because the vast majority of the products fell into the medium masculine-medium feminine group, a second pre-test was conducted.

On the second pre-test a few product changes were made, but the questionnaire was altered considerably. Seventy-eight male and female business students were administered a two-part questionnaire. The first part of the questionnaire asked the respondent to indicate the extent to which he/she felt each product had a masculine image, according to a nine-point horizontal scale with extremes labeled "not at all masculine" (1) and "extremely masculine" (9). The mid-point was labeled "moderately masculine" (5). The second part of the questionnaire elicited identical information for feminine image. Respondents had no difficulty with the second pre-test, and the range of scores was wider. The results, however, indicated that products fell into the same three image categories, but with more products in the high masculine-low feminine and low masculine-high feminine categories. Products were assigned to categories on the basis of a cluster analysis and inspection of mean scores.

A third pre-test on a much smaller sample of students collaborated the results of the second pre-test. On the basis of the pre-test results the study was redesigned to include products in three categories: high masculine-low feminine, medium masculine-medium feminine, and low masculine-high feminine. It is interesting to note that for the 125 products tested there was no such thing as an androgynous (high masculine-high feminine) or neuter (low masculine-low feminine) product.

### Final Questionnaire and Administration

A four part questionnaire was administered to 307 male and female business students. The first part of the questionnaire elicited the masculine image of 24 products representative of the three image categories from the pre-test results. The scale was exactly the same as that of the second and third pre-test. The second part of the questionnaire obtained the feminine

image of the same 24 products.

The third part of the questionnaire consisted of the measure of sex role self-concept. The measure used was Bem's (1974) which consists of 60 adjectives representing extensively tested stereotypical masculine and feminine traits. The respondent indicates on a horizontal seven-point scale the extent to which each adjective describes him or her.

The final section of the questionnaire was a measure of self-esteem (Helmreich and Stapp 1974). This measure has been an effective predictor of behavior in normal situations and tested in a variety of contexts. In addition, it is less related to masculinity than traditional measures, and differences between the sexes are nonsignificant. For these reasons the Texas Social Behavioral Inventory Form B was chosen to measure self-esteem. Finally, respondents' sex was elicited.

### Final Questionnaire and Administration

Respondents were assigned to one of four groups on the basis of their sex role self-concept inventory: androgynous (high masculine-high feminine), masculine (high masculine-low feminine), feminine (low masculine, high feminine), or neuter (low masculine-low feminine). The scoring for Bem's inventory (1974) allows for three categories of sex role self-concept: masculine, feminine, androgynous (where androgynous is defined as equal amounts of masculine and feminine orientation, high or low). However, according to the conceptualization of androgyny offered by Spence, Helmreich and Stapp (1975), this study divided what Bem calls androgynous people into two groups: high identification with traits of both sexes and low identification with traits of both sexes. Thus, the low identification group was termed "neuter". Traditionally, sex role self-concept has been measured on one bi-polar scale. The technique and conceptualization used here (for both persons and products) allows masculinity and femininity to be independent constructs. Bem (1974) provides an excellent discussion of this approach.

For subsequent analyses respondents' self-esteem scores were divided into tertiles. Prior to this breakdown, self-esteem scores were submitted to analysis to determine significant differences by sex, sex role self-concept, and combined sex by sex role self-concept, separately.

Masculine and feminine image scores on the 24 products were used to assign products to one of the three previously determined groups. The groups elicited from the final instrument were identical to the product assignments determined by the pre-tests. T-tests between masculine and feminine mean scores for each product revealed significant differences for 23 of the 24 products. Key ring was the only product for which masculine and feminine image did not differ significantly at the .05 level. Nylon underwear had the highest mean feminine rating (7.90) and the second lowest masculine mean score (2.26). Scarf had a lower masculine rating (1.68) and a high feminine rating (7.85); however, due to previous research for the product (Morris and Cundiff 1971) hair spray is discussed here as the second high feminine (7.38) low masculine (2.79) product. The masculine and feminine means for key ring were 4.42 and 4.50, respectively. It constitutes a moderately feminine, moderately masculine product. Pocket knife had the highest masculine mean (7.71) and the lowest feminine mean (1.86). Cuff links had the second highest masculine mean (7.30) and a low feminine mean (2.74). Thus, pocket knife and cuff links represent high masculine-low feminine products.

Each product was submitted to analysis of variance (Nie, et. al. 1975) for masculine and feminine perceptions



separately. Sex-typed product image constituted the dependent variables and sex role self-concept, sex, use and self-esteem constituted the independent variables which were categorized as previously described. The remainder of the discussion focuses upon the results of these analyses for the five sex-typed products, following the self-esteem analysis.

### Results and Discussion

Significant differences occurred for mean self-esteem scores between the four sex role self-concept groups, ( $\alpha=.000$ ) but not between males and females. The ranking of sex role self-concepts for self-esteem were: androgynous (48.89,  $n=74$ ), masculine (47.20,  $n=80$ ), feminine (40.12,  $n=83$ ), and neuter (37.94,  $n=83$ ). The probability of F was again .000. Combining sex and sex role self-concept yielded the following mean self-esteem scores ( $\alpha=.000$ ): androgynous females (51.51,  $n=31$ ), androgynous males (47.60,  $n=43$ ), masculine males (47.35,  $n=65$ ), masculine females (46.80,  $n=15$ ), feminine females (40.57,  $n=63$ ), neuter females (40.29,  $n=24$ ), feminine males (39.35,  $n=20$ ), and neuter males (36.78,  $n=46$ ). These results are very comparable to those of previous research (Spence et. al 1975) with a reverse ordering of masculine females with masculine males and feminine males with neuter females.

TABLE 1

SUMMARY OF ANALYSIS OF VARIANCE FOR MASCULINE AND FEMININE PERCEPTIONS OF POCKET KNIFE<sup>a</sup>

Effect	Mean Square	df	F	p
Masculine Perceptions				
Sex Role (A)	5.345	3	3.059	.029
Sex (B)	.843	1	.483	.488
Self-esteem (C)	1.414	2	.809	.446
Use (D)	.948	1	.542	.462
A x B	4.718	3	2.700	.046
A x C	3.732	6	2.136	.050
B x C	6.618	2	3.787	.024
B x D	3.676	1	2.104	.148
A x C x D	3.875	5	2.217	.053
B x C x D	4.979	2	2.849	.060
A x B x C x D	11.643	2	6.663	.002
Error	1.747	265		
Feminine Perceptions				
Sex Role (A)	2.509	3	2.748	.043
Sex (B)	9.128	1	9.998	.002
Self-esteem (C)	1.767	2	1.936	.146
Use (D)	4.750	1	5.203	.023
B x D	1.544	1	1.691	.195
C x D	2.459	2	2.693	.069
Error	.913	265		

<sup>a</sup>Interactions with  $p < .20$  are displayed.

Table 1 presents the results of the analysis of variance for masculine and feminine perceptions of pocket knife, and Table 2 exhibits mean scores for sex role, sex and use and significant two-way interactions. In general, the variables tested appear to have a stronger impact on masculine image of the masculine sex-typed product than upon feminine image.

As can be seen from Table 2 for the main effect of sex role upon both masculine and feminine perceptions, feminine persons perceive the product as more masculine than other sex role self-concepts, and androgynous persons perceive the product as most feminine. A

significant main effect of sex and use occurs only for feminine product perceptions. Females and people who use the product perceive it as more feminine than do males and non-users. Self esteem is not an important influence on product perceptions; however, in interaction with other variables, it becomes an important variable for masculine product perceptions.

TABLE 2

MEAN SCORES FOR SIGNIFICANT<sup>a</sup> EFFECTS ON MASCULINE AND FEMININE IMAGE OF POCKET KNIFE

Level	Mean Score		n
	Masculine	Feminine	
Sex Role			
Androgynous	7.676	2.000	74
Masculine	7.638	1.850	80
Feminine	8.121	1.687	83
Neuter	7.357	1.914	70
Sex			
Male	N.S.	1.799	174
Female		1.932	133
Use			
Do not use	N.S.	1.764	127
Use		1.932	180
Sex Role x Sex			
Androgynous Males	7.442	N.S.	43
Masculine Males	7.523		65
Feminine Males	7.350		20
Neuter Males	7.348		46
Androgynous Females	7.484		31
Masculine Females	7.333		15
Feminine Females	7.667		63
Neuter Females	6.792		24
Sex Role x Self-Esteem			
Low SE Androgynous	7.857	N.S.	7
Low SE Masculine	7.571		7
Low SE Feminine	8.000		44
Low SE Neuter	7.683		41
Medium SE Androgynous	7.368		19
Medium SE Masculine	7.675		40
Medium SE Feminine	8.464		28
Medium SE Neuter	7.083		24
High SE Androgynous	7.771		48
High SE Masculine	7.606		33
High SE Feminine	7.727		11
High SE Neuter	6.000		5
Sex x Self Esteem			
Low SE Males	7.679	N.S.	56
Low SE Females	8.023		43
Medium SE Males	7.484		62
Medium SE Females	7.959		49
High SE Males	7.536		56
High SE Females	7.732		41

<sup>a</sup> $p < .05$

The significant interaction of sex and sex role for masculine perceptions indicates that masculine, feminine or neuter females are likely to perceive the product as less masculine than do males of any sex role self-concept. Androgynous females perceive the product as more masculine than do feminine or neuter males.

As Table 2 further indicates, for the significant masculine perception interaction of sex role and self-esteem, medium self-esteem feminines have the most masculine perceptions of the product, and high self-esteem neuters have the lowest. No clear pattern is discernable, however,

For the significant interaction of sex and self-esteem females with low, medium and high self-esteem, respectively, have the highest masculine perceptions. The lowest masculine perception occurs for medium self-esteem males.

Due to cell size difficulties and the resulting instabilities, mean scores are not presented for the significant masculine image interactions of sex role by self-esteem by use or the four-way interaction. It appears that for this product masculine perceptions are influenced by combinations of variables rather than individual variables; thus, these perceptions are not easily explained. There is a slight pattern which tends to indicate that opposites (in terms of sex or sex role) may have stronger masculine perceptions. Feminine perceptions are more straight forward, being influenced by sex role self-concept, sex and use.

The summary of the analysis of variance for the other high masculine-low feminine product, cuff links, is presented in Table 3. Mean scores for significant main effects and two-way interactions are exhibited in Table 4.

Sex role self-concept is only significant for feminine perceptions of cuff links. Unlike the feminine perceptions for pocket knife, the means for sex role self-concept groups for feminine perceptions of cuff links indicate the strongest female perceptions occur for feminine self-concepts and the weakest for neuter self-concept. Females have the strongest sex-typed image for both masculine and feminine perceptions, perceiving the product as both more masculine and more feminine than do males. For masculine perceptions, persons who use the product perceive it as more masculine than do those who do not use cuff links. This was consistent with feminine image results for pocket knife.

TABLE 3

SUMMARY OF ANALYSIS OF VARIANCE FOR MASCULINE AND FEMININE PERCEPTIONS OF CUFF LINKS<sup>a</sup>

Effect	Mean Square	df	F	p
MASCULINE PERCEPTIONS				
Sex Role (A)	3.617	3	1.225	.301
Sex (B)	19.662	1	6.658	.010
Self-Esteem (C)	1.685	2	.571	.566
Use (D)	11.493	1	3.892	.050
A x B	6.258	3	2.119	.098
B x D	26.339	1	8.919	.003
A x C x D	5.118	6	1.733	.114
B x C x D	10.648	2	3.606	.029
Error	2.953	264		
FEMININE PERCEPTIONS				
Sex Role (A)	6.958	3	2.690	.047
Sex (B)	16.054	1	6.206	.013
Self-Esteem (C)	2.088	2	.807	.447
Use (D)	.392	1	.151	.697
A x D	4.563	3	1.764	.154
B x D	20.478	1	7.916	.005
Error	2.587	264		

<sup>a</sup>Interactions with  $p \leq .20$  are displayed.

Again, interactions appear to be more important for masculine perceptions than for feminine perceptions. The interaction of sex and use is significant for both masculine and feminine perceptions for cuff links. It was previously insignificant for both perceptions of

TABLE 4

MEAN SCORES FOR SIGNIFICANT<sup>a</sup> EFFECTS ON MASCULINE AND FEMININE IMAGE OF CUFF LINKS

Level	Mean Score		n
	Masculine	Feminine	
Sex Role			
Androgynous	N.S.	2.960	74
Masculine		2.613	80
Feminine		3.072	83
Neuter		2.257	70
Sex			
Male	7.126	2.480	174
Female	7.526	3.083	133
Use			
Do not use	7.190	N.S.	143
Use	7.400		164
Sex x Use			
Male do not use	6.431	2.863	51
Male Use	7.415	2.317	123
Female Do not Use	7.609	2.815	92
Female Use	7.342	3.683	41

<sup>a</sup> $p \leq .05$

pocket knife. For masculine perceptions, females who do not use the product perceive it as most masculine, followed by males who use, females who use and males who do not use. Males who do not use perceive cuff links as most feminine, followed by females who use, females who do not use, and males who use. These results provide some support for the congruency between sex, product use, and product perceptions.

The three-way interaction of sex, self-esteem and use is significant for masculine perceptions, as it was near significant ( $\alpha = .05$ ) for pocket knife. In order to increase the sizes of some cells a much larger sample is needed (307 used here). Although there are interpretation difficulties, the prevalence of significant and close to significant results for three and four-way interactions on masculine perceptions indicates that this may merit further investigation. The relationships among these variables for masculine perceptions appears complex.

Table 5 and 6, respectively, summarize the analysis of variance and present mean scores for nylon underwear. Sex role is significant only for feminine perceptions with feminine persons perceiving nylon underwear as most feminine and androgynous persons perceiving the product as least feminine. Males perceive the product as more masculine than do females, and although not within an alpha level of .05, females perceive the product as more feminine than do males. Again, people who use the product perceive it as more masculine, but unlike earlier results, people who do not use perceive the product as more feminine. Similar to the masculine sex-typed products, interactions are more important for masculine perceptions of nylon underwear than for feminine perceptions.

For masculine perceptions, the interaction of sex role and self-esteem is significant. As was the case with pocket knife, a pattern is difficult to discern. The highest masculine perceptions occur for low self-esteem androgynous and the lowest occur for low self-esteem masculine.

The significant interaction of sex role and use indicates that masculine, neuter and androgynous persons using the product perceive it the most masculine, with little difference between androgynous people who use and who

don't use. Males who do not use, feminine persons and neuters who do not use perceive the product as least masculine. There is some intuitive congruency between product perceptions, use and sex role self-concept.

TABLE 5

SUMMARY OF ANALYSIS OF VARIANCE FOR MASCULINE AND FEMININE PERCEPTIONS OF NYLON UNDERWEAR<sup>a</sup>

Effect	Mean Square	df	F	p
Masculine Perceptions				
Sex Role (A)	.731	3	.293	.831
Sex (B)	75.541	1	30.255	.001
Self-esteem (C)	2.873	2	1.151	.318
Use (D)	91.113	1	36.492	.001
A x C	5.281	6	2.115	.052
A x D	9.019	3	3.612	.014
B x C	9.837	2	3.940	.021
C x D	15.004	2	6.009	.003
A x B x C	3.876	5	1.553	.174
A x C x D	4.085	5	1.636	.151
Error	2.497	270		
Feminine Perceptions				
Sex Role (A)	5.410	3	2.798	.041
Sex (B)	6.327	1	3.272	.072
Self-esteem	1.079	2	.558	.573
Use (D)	9.324	1	4.822	.029
A x D	3.707	3	1.917	.127
A x B x C	3.130	5	1.619	.155
A x C x D	3.839	5	1.985	.081
Error	1.934	270		

<sup>a</sup> Interactions with  $p \leq .20$  are displayed.

For the interaction of sex and self-esteem, medium and high self-esteem males perceive the product as most masculine, respectively, and low self-esteem males and females perceive it as least masculine. The pattern was different for pocket knife where females had a tendency to see the product as more masculine. The results are consistent, however, that "opposites" see the product as more strongly the opposite. Differences in results may be due to differences in the sex-typing of products.

High self-esteem persons who use nylon underwear perceive the product as most masculine, and low self-esteem people who do not use perceive the product as least masculine. For medium self-esteem persons there is no significant difference in masculine perceptions if they use or don't use the product.

Tables 7 and 8 present the summary of the analysis of variance for hair spray and mean scores, respectively. For the main effect of sex, males perceived hair spray as both more masculine and more feminine. The effect of sex appears to be tied to the product and not necessarily to the sex-type of the product.

Use was not significant for feminine perceptions, but people who use hair spray perceive it as more masculine than persons who do not. In all but one instance (feminine perceptions of nylon underwear) where the effect was significant, people using a product perceived it as having more masculinity or femininity than did people who did not use the product.

Males and females who use hair spray perceive the product as more masculine than do males and females

TABLE 6

MEAN SCORES FOR SIGNIFICANT<sup>a</sup> EFFECTS ON MASCULINE AND FEMININE IMAGE OF NYLON UNDERWEAR

Level	Mean Score		n
	Masculine	Feminine	
	Sex Role		
Androgynous	N.S.	7.647	74
Masculine		7.813	80
Feminine		8.325	83
Neuter		7.771	70
	Sex		
Male	2.431	7.793 <sup>b</sup>	174
Female	2.045	8.045	133
	Use		
Do not Use	2.054	7.925	147
Use	2.456	7.881	160
	Sex Role x Self-Esteem		
Low SE Androgynous	3.286	N.S.	7
Low SE Masculine	1.857		7
Low SE Feminine	1.568		44
Low SE Neuter	2.146		41
Med. SE Androgynous	3.000		19
Med. SE Masculine	2.375		40
Med. SE Feminine	2.179		28
Med. SE Neuter	2.167		24
High SE Androgynous	2.146		48
High SE Masculine	2.727		33
High SE Feminine	2.727		11
High SE Neuter	2.800		5
	Sex Role x Use		
Androgynous Do not use	2.429	N.S.	35
Androgynous Use	2.513		39
Masculine Do not use	2.268		56
Masculine Use	2.958		24
Feminine Do not use	1.667		21
Feminine Use	2.016		62
Neuter Do not use	1.571		35
Neuter Use	2.829		35
	Sex x Self-Esteem		
Low SE Male	2.018	N.S.	56
Low SE Female	1.861		43
Med. SE Male	2.661		62
Med. SE Female	2.041		49
High SE Male	2.589		56
High SE Female	2.244		41
	Self-Esteem x Use		
Low SE Do not use	1.833	N.S.	48
Low SE Use	2.059		51
Med. SE Do not use	2.302		53
Med. SE Use	2.466		58
High SE Do not use	2.000		46
High SE Use	2.843		51

<sup>a</sup>  $p \leq .05$

<sup>b</sup>  $p \leq .072$

who do not use. Males using the product perceive it as most masculine. Males who do not use the product perceive hair spray as most feminine and females who do not use perceive it as least feminine. Again there is indication of a congruency between sex, use and product perceptions.

Hair spray may be a different type of product than the previous three. It is a product which has recently been accepted by more males, although it was sex-typed as high feminine-low masculine. The importance of three-way interactions for feminine perceptions is unique to this product. Although there is some consistency between results for sex, use and sex by use across products, the results for similarity sex-typed products still tends to

be product specific. Sex role self-concept does appear to influence product perceptions and self-esteem appears only to be important in interaction with other variables.

Key ring was the only product for which there was not a perceived difference between masculinity and femininity. This medium masculine-medium feminine, but really "neuter" product is distinctly different from the sex-typed products.

TABLE 7

SUMMARY OF ANALYSIS OF VARIANCE FOR MASCULINE AND FEMININE PERCEPTIONS OF HAIR SPRAY<sup>a</sup>

Effect	Mean Square	df	F	p
Masculine Perceptions				
Sex Role (A)	3.794	3	1.770	.153
Sex (B)	15.406	1	7.188	.008
Self-esteem (C)	3.565	2	1.663	.191
Use (D)	68.827	1	32.111	.001
B x D	12.543	1	5.852	.016
A x B x C	4.315	5	2.013	.077
Error	2.143	262		
Feminine Perceptions				
Sex Role (A)	4.035	3	1.519	.210
Sex (B)	17.886	1	6.735	.010
Self-esteem (C)	1.988	2	.749	.474
Use (D)	.017	1	.007	.936
A x B	4.526	3	1.704	.167
B x D	17.479	1	6.582	.011
A x B x D	8.144	3	3.067	.029
A x C x D	4.432	6	1.669	.129
Error	2.656	262		

<sup>a</sup>Interactions with  $p \leq .20$  are displayed.

TABLE 8

MEAN SCORES FOR SIGNIFICANT<sup>a</sup> EFFECTS ON MASCULINE AND FEMININE IMAGE OF HAIR SPRAY

Level	Mean Score		n
	Masculine	Feminine	
Sex			
Male	2.845	7.529	174
Female	2.707	7.181	133
Use			
Do not use	2.394	N.S.	175
Use	3.303		132
Sex x Use			
Male Do not use	2.407	7.644	118
Male use	3.768	7.286	56
Female Do not use	2.368	6.895	57
Female Use	2.961	7.395	76

<sup>a</sup> $p \leq .20$

Table 9 presents the summary of the analysis of variance for masculine and feminine perceptions of key ring. The only significant variable was self-esteem for feminine perceptions of key ring. Medium esteem persons saw the product as most feminine (4.786), followed by high self-esteem (4.609) and low self-esteem (4.061). This is the only product for which self-esteem was significant as a main effect. If generalizations can be made from one product, this data indicates that for non-sex-typed products (which according to the pre-test and results are very rare) sex, sex role self-concept, and product use have little influence on masculine and feminine product perceptions. This result is intuitively appealing since the product is not sex-typed these variables would not be expected to be extremely important.

Conclusions and Implications

Several conclusions can be drawn from the results of this research. First, products have very definite masculine and feminine sex-typed image. However, the sex-typing of people and products is very different. Although it is possible to have masculine, feminine, androgynous and neuter people without forcing an analysis of extremes, this apparently is not the case with products. For products, sexual image is still a bi-polar construct for strong masculine and feminine identity, or an equally weak one. Further, masculinity and femininity of products appears to be related to which sex is most often thought of as using the product. This conclusion does not emerge clearly from the data provided here; however, it is an implication worthy of further investigation. If product images are derived from the sex perceived as using the product most frequently (a suggested methodological approach), then the marketer interested in changing product perceptions should change the sex featured as using the product. Hair spray may be an example of a product in the process of change. There may be learning involved in product sex-typing. And, sex role self-concept does appear to be an important influence on the sex typing of products.

TABLE 9

SUMMARY OF ANALYSIS OF VARIANCE FOR MASCULINE AND FEMININE PERCEPTIONS OF KEY RING<sup>a</sup>

Effect	Masculine Perceptions			
	Sex Role (A)	2.656	3	.694
Sex (B)	.266	1	.070	.792
Self-esteem (C)	.930	2	.243	.785
Use (D)	4.743	1	1.239	.267
A x C x D	9.305	1	2.430	.120
B x C x D	11.300	1	2.951	.087
Error	3.829	275		
Effect	Feminine Perceptions			
	Sex Role (A)	3.003	3	.982
Sex (B)	.216	1	.071	.791
Self-esteem (C)	10.684	2	3.494	.032
Use (D)	.821	1	.269	.605
A x C x D	6.003	1	1.963	.162
Error	3.058	275		

<sup>a</sup>Interactions with  $p \leq .20$  are displayed.

In addition, there appears to be something different operating among the variables sex, sex role self-concept, use and self-esteem for masculine and feminine perceptions. Across sex-typed products, feminine perceptions appear to be more related to individual variables than combinations of variables. The opposite is true for masculine perceptions which almost seem more complex. However, sex, sex role self-concept, use and self-esteem do, individually and collectively, influence product image. For sex-typed products, self-esteem is only important in conjunction with other variables.

Although there are consistencies in the variables which were important across products, directionality appears to be product specific. There is limited support for the conclusion (cuff links and hair spray) that males will perceive a feminine-typed product as more sex-typed for both masculine and feminine perceptions and females will perceive a masculine-typed product as more sex-typed for both masculine and feminine product image. Also, the more extremely typed the products (e.g. pocket knife and nylon underwear), the more the variables tested appear to influence product image.

A strong, systematic relationship among product perceptions and sex, sex role self-concept, use and self-

esteem across products could not be identified from the results of this study. It can be concluded, however, that the variables investigated do impact upon masculine and feminine product perceptions. Further investigation into the nature of the relationship between these variables, operationalized in other contexts, may provide the marketer with another perspective on the consumer behavior impact of changing sex roles for women--and, more importantly, for both sexes.

#### References

- Bem, Sandra L. (1974) "The Measurement of Psychological Androgyny," Journal of Consulting and Clinical Psychology, 42, 155-62.
- \_\_\_\_\_ (1972), "Psychology Looks at Sex Roles: Where Have All the Androgynous People Gone?" Presented at the UCLA Symposium on Women.
- Bielliaskas, Vytautas, and Mikesell, Richard H. (1972), "Masculinity-Femininity and Self-Concept," Perceptual and Motor Skills, 34, 163-67.
- Birdwell, A.E. (1968), "A Study of the Influence of Image Congruence on Consumer Choice," Journal of Business, 41, 76-88.
- Deutsch, Connie J., and Gilbert, Lucia A. (1976), "Sex Role Stereotypes: Effect on Perceptions of Self and Others and on Personal Adjustment," Journal of Counseling Psychology, 23, 373-79.
- Dolich, I.J. (1969), "Congruence Relationships Between Self Images and Product Brands," Journal of Marketing Research, 6, 80-84.
- Feather, N.T., and Raphelson, Alfred C. (1973), "Fear of Success in Australian and American Student Groups: Motive or Sex-Role Stereotype?" Journal of Personality, 42, 190-201.
- Fry, Joseph (1971), "Personality Variables and Cigarette Brand Choice," Journal of Marketing Research, 8, 298-304.
- Gentry, James, Doering, Mildred and O'Brien, Terrence (1978), "Masculinity and Femininity Factors in Products Perceptions and Self Image," Advances in Consumer Research, 5, 326-332.
- Gorden, Francine E., and Hall, Douglas T. (1973), "Self-Image and Stereotypes of Femininity: Their Relationship to Women's Role Conflicts and Coping," Journal of Applied Psychology, 59, 241-43.
- Grubb, Edward L., and Hupp, Gregg (1968), "Perceptions of Self, Generalized Stereotypes, and Brand Selection," Journal of Marketing Research, 5, 58-63.
- \_\_\_\_\_, and Grothwohl, Harrison L., "Consumer Self-Concept, Symbolism and Market Behavior: A Theoretical Approach," Journal of Marketing, 31, 22-27.
- Hamm, B.C., and Cundiff, Edward W. (1969), "Self-Actualization and Product Perceptions," Journal of Marketing Research, 6, 470-72.
- Helmreich, Robert, and Stapp, Joy (1974), "Short Forms of the Texas Social Behavior Inventory (TSBI), an Objective Measure of Self-esteem," Bulletin of the Psychonomic Society, 4, 473-75.
- Hughes, G. David, and Guerrero, Jose L. (1971), "Automobile Self-Congruity Models Reexamined," Journal of Marketing Research, 8, 125-7.
- Jordan-Viola, Eunice, Fassberg, Susan, and Viola, Michael T. (1976), "Feminism, Androgyny, and Anxiety," Journal of Consulting and Clinical Psychology, 44, 870-71.
- Kravetz, Diane (1976), "Sex Role Concepts of Women," Journal of Consulting and Clinical Psychology, 44, 437-443.
- Landon, E. Laird, Jr. (1974), "Self Concept, Ideal Self Concept and Consumer Purchase Intentions," Journal of Consumer Research, 1, 44-51.
- Levy, S.J. (1959), "Symbols for Sale," Harvard Business Review, 37, 117-24.
- Morris, George, and Cundiff, Ed (1971), "Acceptance by Males of a Feminine Product," Journal of Marketing Research, 7, 372-374.
- Nie, Norman H., Hull, C. Hadlai, Jenkins, Jean G., Steinbrenner, Karin, and Brent, Dale H. (1975), Statistical Package for the Social Sciences, McGraw-Hill: New York.
- Seyfried, B.A. and Hendrick, Clyde (1973), "When Do Opposites Attract? When They are Opposite in Sex and Sex-Role Attitudes," Journal of Personality and Social Psychology, 25, 15-20.
- Shaffer, David, R., and Wegley, Carol (1973), "Success Orientation and Sex-Role Congruence as Determinants of the Attractiveness of Competent Women," Journal of Personality, 42, 585-600.
- Spence, Janet T., Helmreich, Robert, and Stapp Joy (1975), "Ratings of Self and Peers on Sex Role Attributes and Their Relation to Self-Esteem and Conceptions of Masculinity and Femininity," Journal of Personality and Social Psychology, 32, 29-39.
- Strober, Myra H., and Weinberg, Charles B. (1977), "Working Wives and Major Family Expenditures," Journal of Consumer Research, 4, 141-47.
- Thorne, Brandon (1977), "The Measurement of Femininity," Journal of Clinical Psychology, 33, 5-10.
- Tolor, Alexander, Kelley, Bryan R. and Stebbins, Charles A. (1976), "Assertiveness, Sex-Role Stereotyping, and Self-Concept," Journal of Psychology, 93, 157-64.

UNDERSTANDING AND ANALYZING CONTINGENCY DATA

Edgar A. Pessemier  
 Krannert Graduate School of Management  
 Purdue University

When analyzing individuals' manifest judgments about the attributes of objects, the process by which these data are generated becomes a subject of real interest. Although the object's attributes may be continuous measurable or discrete, in principle the individual's response to any object-attribute stimulus must be categorical (Young). Not infrequently, the response categories are very limited. For example, a judge can say yes, no, or don't know to the question, "Does this object have this attribute?" Frequently, these responses will be code 1 if yes and otherwise 0. Other coding schemes may be used, e.g., yes = 1, don't know = 0 and no = -1. Several methods for analyzing this type of perceptual data will be described at a later point. Now, attention will be devoted to examining alternative psychological processes which might lead to individual judgments in the form noted above.

A large experimental literature supports the general concept of discriminial dispersion first described by Thurstone (59) and recently elaborated by Luce (77), Yellott (77) and others. This theory states that an external stimulus or signal generates an internal representation. The transformation is stochastic, producing different internal values in response to each exposure to the stimulus. These internal representations form a distribution along a psychological continuum. This distribution can be characterized by its location, shape and dispersion. In most work, a normal, lognormal or logit distribution has been employed. When the particular theory is applied to comparative and categorical judgments, it is possible to scale stimuli and/or stimuli boundaries along a psychological continuum.

I. Comparative Judgments (See Bock & Jones 68)

Here, a standard stimulus is selected and each of the remaining stimuli are compared to the standard. Each of the remaining stimuli are judged to be greater than, equal to or less than the standard stimulus. The instantaneous internal values of the standard and the remaining stimuli determine whether each stimulus is judged to be greater than, equal to or less than the standard. In turn, these instantaneous values depend on the location, shape and dispersion of the standard and other stimuli.

In the present context, the standard stimuli can be thought of as a threshold level for responding yes to the question, "Does this object have this attribute?". If yes, don't know, and no responses are admissible, then a high and low standard can be used. When the instantaneous internal value of the test stimulus falls above the instantaneous value of the high threshold, the response is yes. When the test stimulus value falls below the low threshold, the response is no. When the test stimulus value falls between the high and low thresholds, the response is don't know.

In the above discussion, responses are assumed to have purely subjective, perceptual content. If objectively correct responses can be determined, there is no clear cut payoff for making the right judgment. Furthermore, in the case of many object-attribute stimuli, e.g., "Does car A have a racy appearance?", an objectively correct response cannot be specified. In these cases, judgmental thresholds will tend to be subject specific. Some subjects will provide more yes, don't know, or no

responses than other subjects but their response category propensities on a judgment will not be associated with the payoff from a correct or in correct choice of category. Under these conditions there appears to be little reason for response criterion to reflect anything but the internal value of the stimuli and past experience with encoding the value as a yes, don't know and no. Nevertheless, the attribute may influence the criterion. A difficult-to-judge attribute (larger dispersion) may spread the locations of the thresholds, increasing the proportion of don't know judgments.

The thresholds' locations (and spread) may be selected in relation to experience such that all objects are judged to have some significant proportion correctly judged in the yes and/or no categories. In the absence of these results, the threshold would serve no useful purpose. The conflicting forces determining the best proportion relate to discrimination and information content. Rarely observed values will contribute little to the average capacity to discriminate among objects but rare observations are strong discriminators of single objects [16]. An individual's selections of response thresholds may be subconscious and inaccessible to direct observation. Furthermore, they may be subject to both unconscious learning and maturation.

II. Signal Detection Theory (See Coombs, Dawes & Tversky 70)

Signal detection theory separates the sensory aspects of how an individual encodes each stimulus and how an individual makes decisions about responses to the internal value of the stimulus. Here, the signal being sent is known. Due to noise in transmission or encoding error, the internal values of signal or stimuli may not be a good representation of the known values. Furthermore, the response of an individual may be conditioned by the payoffs associated with giving a particular response, when this response does or does not match the stimulus or transmitted signal. Take a simple case where y = signal sent, n = no signal and a respondent must answer Y = yes, signal present, or N = no, signal absent. The possible signal-response states are:

$$V_{Yy}, - V_{Ny}, - V_{Yn}, V_{Nn}$$

Since the V's are the values assigned to outcomes, the expected value of a yes and no, given an observation x (and its associated internal value) is

$$E(Y|x) = V_{Yy}p(s|x) - V_{Yn}p(n|x)$$

and

$$E(N|x) = V_{Nn}p(n|x) - V_{Ny}p(s|x).$$

The response should be yes as long as  $E(Y|x) \geq E(N|x)$ . As noted below, the best decision is the one for which the response is yes if the likelihood ratio,  $l(x)$ , is greater than or equal to the threshold,

$$l(x) = \frac{p(x, y)}{p(x, n)} \geq \frac{p(n)}{p(y)} \cdot \left[ \frac{V_{Nn} + Y_{Yn}}{V_{Yy} + V_{Ny}} \right] = \beta.$$

The likelihood ratio      Prior odds      Prior costs      Yes response threshold

In the context initially described here, judgments are assumed to be perceptual and outcome costs are very and/or roughly equal. When this is the case,

$$\beta = \frac{p(n)}{p(y)}, \text{ so an individual says yes when}$$

$$\frac{p(x, y)}{p(x, n)} \geq \frac{p(n)}{p(y)} = \rho.$$

The subjective probabilities on the left hand side of the above equation are functions of the individual's discriminational processes.

### III. Test Theory (Birnhum 68, Rasch 66).

The general concepts and computation apparatus employed in theories of comparative judgments and in signal detection theory can be used in more complex categorical judgment tasks. In the case at hand, these extensions will contribute little new. Test theory, however, offers some interesting additional insight, particularly as it relates to the Law of Comparative Judgments (Brogden, 77). The Rasch and Birnhum models express the probability that individual  $i$  will correctly respond to item (signal or stimulus)  $a$ , as:

$$P_{ia} = \frac{\exp(\theta_i - \delta_a)}{1 + \exp(\theta_i - \delta_a)},$$

where

$\exp \theta_i$  = the individual (ability) parameter,

and

$\exp \delta_a$  = the stimulus (difficulty) parameter.

The model involves distributions of ability across individuals and distributions of difficulty across tests. Assuming the item's difficulty and the respondents ability all refer to the same domain, a common scale may be employed. Therefore, responses are the result of sampling from the two distributions and computing  $p_{ia}$  as a function of the distance  $\theta_i - \delta_a$ . Note that

$$\theta_i - \delta_a = L^{-1}(p_{ia}) = \ln \left[ \frac{p_{ia}}{1 - p_{ia}} \right],$$

where  $L^{-1}(p_{ia})$  is the logistic inverse, and is similar to the normal inverse that appears in the Law of Comparative Judgment.

### IV. Other Related Models

Closely related to the test theory model is the Guttman scale (Torgerson, 58) along which both stimuli (items) and individual respondents may be scaled. In this case, it is assumed that items can be located along a continuum according to the degree to which they possess the attribute being scaled and that individuals can also be located along the same continuum. Items above the respondent on the scale will receive one response (yes, no; agree-disagree;...) and items below the respondent will receive the other response. Since this model is deterministic, the location of the stimuli and individual forecast the individual's complete response set.

Other related procedures are various kinds of bioassay and reliability theory models where the strength of treatment or time in use is used to predict the proba-

bility of an event such as death, recovery, purchase, failure etc. Events are discrete and the treatments increase the probability that an event occurs. Individuals have different thresholds at which the treatment level or stimulus triggers the event. The use of Probit, Normit or similar form of analysis lead to a linear equation linking the amount of the treatment with the predicted proportion of the treated population that will respond, e.g., die, recover, buy the product once, or say yes. When this is the case, alterations in the physical properties of products may be used to forecast the proportion of all potential buyers who would respond that the product had a particular attribute, say sweet flavor.

Finally, multivariate versions of some of the above models may also be found. For example, the multivariate logistic distribution might be used to estimate the degree to which two or more variables influenced the proportion of the treated subjects who positively responded.

### Summary

To one degree or another, the above models represent response theories linking specific types of responses, yes, no, and possibly don't know to the question, Does this stimulus object have this attribute? Each model includes specific assumptions about human perceptual judgments and responses but differences are less noticeable than similarities. All models involve an internalized value of each stimulus and one or more thresholds or standards used by subjects to make categorical responses. None of the theories directly indicate how to determine the appropriate response for an individual since the internal value of the stimulus is not observable and the consequences of Y, N, DK responses to purely perceptual judgments offers little guidance. Correctness, truthfulness, and social conformity may be significant motives but it is far from clear how to incorporate these factors in models of individual response.

### Information and Discrimination

Having devoted attention to some alternative response models for contingency data, it is useful to look at the general behavioral influences affecting the observed responses in a particular case. Three elements deserve special attention; how objects enter an individual's evoked set, the attributes that are employed in the perceptual process applied to this set and how objects are encoded into contingency data.

The Evoked Set. For any class of objects, the more similar a new object is to the objects already in an individual's evoked set, the less perceptual variety it will contribute to an evoked set. Generally, the utility of perceptual redundancy among objects is low. It is the different object that gets attention and becomes a new addition to the set of objects about which the individual is most aware. Therefore, individual search is directed towards finding collections of objects that differ in important ways from each other and not simply adding highly similar objects.

The Attribute Set. Given the above incentives, it is plausible to believe individuals will use attributes that effectively discriminate among the objects in their evoked sets. Non-discriminating attributes do not help an individual identify genuinely new or surprising objects. Therefore, individuals are expected to store and use attribute data that help enhance the useful variety of objects in their evoked sets. Their attributes are learned from experience and usually have evaluative content.

The Stored Contingency Data. At this point it is useful to recall that the information content in a message, Object \_\_\_\_\_ has (does not have) attribute \_\_\_\_\_, depends on the probability that any object of the type in question has the attribute. The information content of an event whose probability of occurrence is  $x$  is

$$h(x) = -\log x.$$

The expected information (before knowing which object is being evaluated) is

$$H(x) = -\sum_{i=1}^n x_i \log x_i$$

and the expected information is a maximum when  $x_i = .5$  for each event (Theil, 67). Therefore, before knowing which object in a set must be identified, an individual will select independent attributes such that each one is possessed by just half the objects and there is no uncertainty (don't know judgments) about the presence or absence of the attribute.

#### Summary

Individuals seek evoked sets of objects (objects about which they are knowledgeable) which are varied and can be readily identified. To accomplish this purpose, the individual selects meaningful attributes across which objects can be expected to significantly vary. Finally, the selection of attributes, objects and the encoding scheme seeks to balance the presence and absence of each attribute across each object, thereby maximizing the expected information content of messages about the set's perceptual content.

#### Information, Learning and Maturation

When confronted with questions about the presence or absence of an attribute for a genuinely new or novel object, the candid individual must answer "don't know" to every object attribute question. As learning takes place, an assortment of determinant attributes and growing number of non-redundant objects are associated with the evoked set. In this sense, the perceptual map of the individual grows in both complexity and clarity. Early, each new stimulus tends to have a high degree of novelty and receives much attention. Later, additions to the "new" class of objects tend to be fitted into an established structure and it becomes progressively more difficult to make the individual aware of new additions. At some point, knowledge about the evoked set may become unitized in the manner described by (Barbara Hayes-Roth, 77) and the objects and their attribute evaluations may become resistant to change. Ultimately, new additions will tend to displace old objects in the evoked set.

A second way to look at the learning process is to re-define (Luce's 77)  $\beta$  operator non-linear learning model. To do so, let the strength of a response to object  $i$  and attribute  $j$  on trial  $n$  be  $v_{ijn}$ . The strength can be a simple inverse function of the discriminant dispersion of object  $i$  along the  $j^{\text{th}}$  attribute continuum. The receipt of an informative exposure on trial  $n$  will increase  $v_{ij,n-1}$ , the strength of the response (or decrease the prior dispersion of  $i$  along  $j$ ). The constant proportional change from trial to trial is  $a_{ij}$ . If no informative exposure is received on trial  $n$ , the prior strength  $v_{ij,n-1}$  will decrease by the constant proportional change  $b_{ij}$ . Therefore, one of two events happens on each trial,

$$v_{ijn} = a_{ij} v_{ij,n-1} \quad \text{informative trial}$$

or

$$v_{ijn} = b_{ij} v_{ij,n-1} \quad \text{uniformative trial}$$

where

$$v_{ijn} = (\text{discriminal dispersion of } i \text{ on } j \text{ at trial } n)^{-1}$$

and

$$a_{ij} > 1 > b_{ij}.$$

For a simple, single threshold model, it follows that the probability of a yes on the  $n^{\text{th}}$  trial is

$$P_{ijn} = \frac{\beta_{ij} v_{ij,n-1}}{1 + \beta_{ij} v_{ij,n-1}}$$

when

$$\beta_{ij} = a_{ij}/b_{ij}.$$

The time series profile of various learning and forgetting sequences can be represented by the above model and it can be adapted to cover yes, no and don't know responses.

#### Preliminary Data Analysis

Once a set of contingency data has been collected from a group of individuals, two questions arise; how, if at all, can the data be pooled and what, if any, preliminary transformations may be useful? The first question largely concerns the degree to which response differences are due to noise or due to heterogeneous classes of respondents. Some types of individuals may be well informed and perceptive while others may be ignorant and unperceptive. Whatever factors may be at work, an initial effort to find response typologies by clustering methods should be instructive. It may indicate the absence of important response taxonomies or it may lead to classifying the subjects according to their response profiles. In the latter case, each group's contingency data would be separately analyzed and reasons for group difference could be investigated by discriminant analysis or analysis of variance methods.

The second data question concerns potential data transformations. Several important related issues will be discussed in later sections. Here, three transformations deserve comment. If differences in the overall frequencies of yes, no and don't know responses are present, normalizing the values of the responses may be considered. Doing so changes the numerical values of yes (and possible don't know and no) responses so each subject's data sum equals the mean sum across all subjects. Another possible transformation would weight each subjects object-attribute response by the information content of the object-attribute responses for the average subject. In other words, more weight would be assigned to rare attributes, those possessed by only a few objects. Third, a priori weights could be assigned to responses for each object and/or attribute to achieve some research objective. For example, effective salience weights might be employed.

Whatever analysis sample or samples are finally chosen, and/or whatever preliminary transformation has been applied, one or more aggregate data matrices are employed in subsequent analyses. Each cell of each matrix contains the sum across subjects of the entries in an individual's object by attribute contingency matrix. Finally, if repeated measures are obtained over time, one or more aggregate matrices would represent each cross section.



Another way to look at the yes, no, don't know response to perceptions about objects is to consider a multi-dimensional space spanned by components of object attributes. For an extended discussion of such spacial models, see (Green and Carroll, 77), (Lingoes 77) and (Pessemier 77), and the associated bibliographies. Consider a perceptual space (component reduced space) containing 3 objects, x, y and z, and three (unit length) attribute vectors. Here, orthogonal projections of objects on attributes defined the attribute levels of each object.

The above perceptual structure can be used to generate contingency data of the type discussed above by observing that

- a) Objects with high (positive) attribute levels will be judged "attribute present" and objects with low (negative) attribute levels will be judged "attribute absent." The present-absent region may be the space falling outside a hypersphere with its center at the origin of the space. Alternatively, it may be the regions beyond asymmetric cut offs along each attribute.
- b) DK responses will occur whenever the internal value of an object falls inside the boundaries of the present-absent region. The effect of this region on DK responses is similar to the effect of discriminial dispersion on DK responses. When discriminial dispersions are large, the don't know response rate is relatively high. This rate decreases as discriminial dispersion decreases. When the hypersphere is large, more values fall in the don't know response region and as it decreases, the rate of don't know responses declines.

In the above formulation, early in the life cycle of a product or product class, products will tend to have a large diameter hypersphere and high don't know response rates. Later in the product life cycle, the diameter of the hypersphere will shrink to reflect the increasingly clear product perceptions (small discriminial dispersions of more experienced individual judges). This time-dependent change is likely to be a significant aspect of the adoption-diffusion process. Advertising and the observation of others doubtless rank high among influences on the rate at which the perceptual process clarifies.

Although the details will not be discussed here, simulated perceptual data can be readily generated which reflect the above structural characteristics and dynamic behavior. The simulated data are initially produced in interval form and then converted to yes, no, don't know responses. Therefore, the simple recovery characteristic of various multivariate procedures can be directly examined. In this manner, changes in discrimination and recovery can be examined over time and for data that reflect various levels of perceptual clarity. Finally, the reduced perceptual spaces produced with metric data can be compared to perceptual spaces developed from the analysis of contingency data. Several approaches are discussed in the next section which assume a homogenous population.

#### Models for the Analysis of Contingency Data

The simplest approach to the analysis of contingency data would be non-dimensional. For example, the THAID algorithm could find the extent to which individuals' yes, no and don't know judgments about the presence or absence of attributes discriminate one object from another (Morgan and Messenger, 73). The result of such an analysis would be a hierarchical scheme or tree

which indicated how attribute judgments could be best used to identify an object. When developed at regular time intervals, the trees could easily represent the perceptual-communications dynamics of a product class over the adoption diffusion-cycle or the product life cycle.

A second approach is dimensional. It uses object-attribute contingency data solely for the purpose of developing a proximity matrix from pooled contingent judgments. A number of proximity measures, e.g., weighted or unweighted distances and correlation indices, could be used for this purpose. Depending on the choice of proximity measure and the specific objective, a variety of metric and non-metric multidimensional scaling algorithms are available to obtain a reduced space object configuration. The inter-object distances in this space can be interpreted as object-object dissimilarities. As an aid to interpretation, attribute (vectors) may be located in the same space.

A third approach transforms proportions or probabilities. Thurstone's theories of discrimination suggest the use of response function deviates. If  $p_{ij}$  is the proportion of yes answers to the question, "Does object i have attribute j?" and  $\hat{g}^{-1}$  is the inverse response function, the desired deviate (distance) measure is

$$y_{ij} = \hat{g}^{-1}(p_{ij}).$$

The  $y_{ij}$  values would appear in each cell of the attribute by object matrix which is subjected to further analysis.

A related approach transforms the proportion or probability into the entropy of the observation (message). In this case, the cell entries in the transformed contingency matrix are

$$y_{ij} = \ln(p_{ij}).$$

With respect to either method outlined above, the question of how proportions are computed becomes a question of concern. For example, it may be computed over all possible positive responses, over all possible positive responses less the don't know responses, or over only the positive responses for an attribute. The following section will illustrate some of these possibilities.

A fourth approach is also dimensional but makes direct use of the pooled contingency judgment. Ries [14] has described the basic computations required to use the approach originally suggested by Benzecri [1]. The model described below expands their work by including yes, no and don't know judgments. Subjects judge  $m$  objects,  $O_m$ , by noting the presence or absence of each of  $N$  primary attributes,  $A_n$ . The presence of an attribute is recorded by a one, the absence of an attribute is recorded by a zero and uncertainty about the presence or absence of an attribute is recorded by a blank. By using an "ambiguity" attribute to record the number of uncertain attribute judgments or blanks, the level of perceptual ambiguity surrounding each object can be easily measured.

Recording as many positive attributes as necessary to describe an object extends the perceptual domain used by (Ring 78) where only one object, the one judged to have the most (or least) of an attribute, received a score of one. Furthermore, including an ambiguity attribute and allowing uncertainty judgments insures that an explicit judgment is recorded for every attribute on every object. The judgments made by single subjects can be recorded in a matrix format. Pooled data for groups of subjects can be recorded in a similar matrix layout. If attributes are rows, for any object column, the count of 1, measures the object's

attribute complexity, the count of Os measures the attribute's simplicity, and the count of blanks measures the object's ambiguity. On the other hand, attributes can be analyzed by rows where the count of Is measures the attribute's rarity and the count of blanks measures the attribute's ambiguity. If the row, column and grand mean of non-zero entries in the data matrix are  $n_{i.}$ ,  $n_{.j}$ , and  $n_{..}$ , then a matrix of "normalized deviations" has entries given by

$$y_{ij} = \frac{n_{ij} - \frac{n_{i.} n_{.j}}{n_{..}}}{\sqrt{\frac{n_{i.} n_{.j}}{n_{..}}}}$$

The Y matrix of "normal deviations" is decomposed and plotted by the computational procedure described by (Ries, 74) and programmed for convenient application by (Jarboe 78). The output of this analysis yields a product map in which the attribute "product ambiguity" is appropriately represented and onto which the level of each product or brand can be projected.

The static characteristics of this map are of interest and can be used for a variety of analytical purposes. Of greater interest, however, are dynamic characteristics of the map, particularly during the earlier phase of a product's or product class' life cycle. During this period, the repeated use of contingency maps of the type just described can track the changing levels of primary product attribute perceptions and the changing levels of ambiguity about products and their various attributes.

#### Summary

The practical value of contingency data has been emphasized. Investigators who need simple response modes, e.g., in telephone surveys, yes-no or yes-no-don't know responses may be essential. Also, respondents usually find these categories are natural verbal responses to a wide range of judgmental questions. In questionnaires, these formats are among the easiest for respondents to use and for investigators to code for further analysis.

The main section of the paper briefly described alternative psychological models for generating individual judgments recorded in the forms noted above. Principal attention was devoted to judgments about continuous attributes of objects. These data are needed to produce perceptual maps for various classes of choice objects such as candidates, automobiles or restaurants. Also, interpretation of the frequency of positive and don't know responses was discussed, particularly as the rate may be a function of individual information processing and maturation.

The final section dealt with various procedures which can be used to transform contingency data and produce spatial representations of objects that are interpretable in terms of determinant attributes. The direct analysis of don't know responses was also discussed. Preliminary field tests of several methods of data collection, reduction and interpretation are underway. Perhaps other investigators will test the appropriateness of the alternative psychological models which may underlie the manifest data.

#### References

- J. P. Benzecri, L'analyse de donnees Vol. 2, L'analyse des correspondences, Dunod, Paris, 1973.
- A. Birnbaum, "Some Latent Trait Models and Their Use in Inferring an Examinee's Ability," in F. M. Lord and R. N. Novick, eds., Statistical Theories of Mental Test Scores, Addison Wesley, 1968.
- R. Bock and Lyle V. Jones, The Measurement and Prediction of Judgment and Choice, Holden-Day, 1968.
- H. E. Brogden, "The Rasch Model, The Law of Comparative Judgment and Additive Conjoint Measurement," Psychometrika, Vol. 42, December 1977, pp. 631-634.
- Clyde H. Coombs, Robyn Dawes and Amos Tversky, Mathematical Psychology, Prentice-Hall, 1970, pp. 165-200 and 270-273.
- Paul E. Green and J. Douglas Carroll, Mathematical Tools for Applied Multivariate Analysis, Academic Press, 1976.
- Glen R. Jarboe, Market Map Analysis of Patronage Behavior, unpublished dissertation, Purdue University, (in preparation).
- Barbara Hayes-Roth, "Evolution of Structures and Processes," Psychological Review, 94, May 1977, pp. 260-278.
- James C. Lingoes, ed., Geometric Representations of Relational Data, Ann Arbor: Mathesis Press, 1977.
- Duncan R. Luce, "Thurstone's Discriminant Process Fifty Years Later," Psychometrika, Vol. 42, December 1977, pp. 461-489.
- James N. Morgan and Robert C. Messenger, THAID: A Sequential Analysis Program for the Analyses of Nominal Scale Dependent Variables, Institute for Social Research, University of Michigan, 1973.
- Edgar A. Pessemier, Product Management: Strategy and Organization, Wiley/Hamilton, 1977, pp. 205-258.
- G. Rasch, "An Individualistic Approach to Item Analyses," In P. L. Lazarsfeld and N. W. Henry, ed., Readings in Mathematical Social Science, Chicago: Science Research Associates, 1966.
- Paul N. Ries, "Joint Space Analysis of Contingency Data," Proctor & Gamble, 1974.
- Larry J. Ring and Charles W. King, "A Multiple Discriminant Analyses Approach to the Development of Retail Store Positioning," Advances in Consumer Research, H. Keith Hunt, ed., Vol. V, Ann Arbor: Association for Consumer Research, 1978, pp. 227-234.
- Henri Thiel, Economics and Information Theory, North-Holland, 1967.
- L. L. Thurstone, The Measurement of Value, University of Chicago Press, 1959.
- Warren S. Torgerson, Theory and Methods of Scaling, John Wiley & Sons, 1958.
- John I. Yellott, Jr., "The Relationship Between Luce's Choice Axiom, Thurstone's Theory of Comparative Judgment, and the Double Exponential Distribution," Journal of Mathematical Psychology, 15, 109-144 (1977).
- Forrest W. Young, "Optimal Scaling With a Variety of Models," Psychometric Laboratory, University of North Carolina, (undated).

Zarrel V. Lambert, Auburn University

## Introduction

'Tis an appropriate season to be a discussant this time of year because the role is akin to that of a Monday morning quarterback. One can say the team should have passed more or a different research design should have been employed. A discussant, somewhat like a Monday morning quarterback, does not have to make hard decisions before the data are collected, analyzed, and problems perhaps become apparent. In addition to having hindsight, a discussant may raise conceptual and methodological questions which the researchers had difficulty answering in advance due to length restrictions on their papers. A discussant's role, like a critic's, is a comfortable one.

More seriously, a positive and productive purpose in discussing research papers is to offer observations and raise questions, perhaps both conceptual and methodological, that may be helpful to future investigators who read the papers prior to formulating their research. My remarks are made with this purpose in mind, and it is hoped they will not be interpreted in such fruitless terms as praise and criticism.

Dissimilarities of the three papers seemed more apparent than commonalities. Those by Golden et al. and Hirschman et al., of course, were data based, whereas Pessemier's was mainly conceptual. The principal perspective of the major research questions appeared to differ between the two empirical papers. The work by Golden et al. seemed to emanate primarily from conceptual questions with applied or managerial implications less in the forefront. By comparison, the research by Hirschman et al. appeared centered largely around managerial questions pertaining to segmentation of credit card users, with conceptual implications more in the background. Because this characterization of the papers is a gross simplification, the authors and others have complete liberty to disagree.

Because the papers seemed to be quite different, they will be discussed individually, instead of collectively, in alphabetical order of the first author's name.

## Sex Role Self-Concept by Golden et al.

Factors in general that influence product perceptions obviously have considerable conceptual and managerial significance. Those pertaining to sex in particular are of current interest and may remain topical for some time if there is continued emphasis on breaking down traditional sex stereotypes. One of this paper's more important contributions is to introduce the possibility that product perceptions may be associated with a person's sex role self-concept.

The paper focuses on reporting statistical findings and devotes scant attention to discussing two key concepts: (1) sex role self-concept; and (2) treatment of femininity and masculinity as separate independent dimensions. This may leave some readers who are unfamiliar with these concepts with unresolved questions in their minds. Some readers, for example, who envision femininity/masculinity as bi-polar anchors on a single dimension might begin arguing with the measurements obtained by the study's dual scales before turning to the cited references pertaining to this concept.

The emphasis on statistical results also may cause some readers difficulty in relating these findings back to the major research assumptions and objectives. The introductory paragraphs seem to indicate that cause-and-effect relationships are not being implied between sex role self-concept, self-esteem, and feminine/masculine product perceptions. However, several of the stated research objectives might be interpreted as suggesting causal relationships. For example, the second objective was: "To determine the individual and combined effects of sex, use and self-esteem upon sex-typing of products." Little attention is given to explaining whether the statistical results reveal associations or cause-and-effect relationships. If it is the latter, then supporting conceptual and/or methodological arguments would be helpful.

Also there was some difficulty in my case in locating direct test results for the first research objective which was "to investigate the congruency between sex role self-concept and sex-typing of products."

The authors, given their knowledge about the various concepts involved in the research, could assist readers by discussing conceptual and/or managerial implications of the statistical results. Otherwise readers who may be less familiar with the concepts might overlook significant implications. For instance, what do the observed differences in self-esteem across the four sex role self-concept classifications mean? Other questions pertaining to implications of the findings for pocket knife, the first product examined, include:

1. Why do feminine product perceptions differ between males and females, and masculine perceptions do not? What are the conceptual and/or managerial implications?
2. Similarly why do feminine perceptions vary between users and nonusers, and masculine perceptions do not? Why would users of pocket knives see the product as being more feminine than nonusers?
3. What inferences can be drawn from the statistical results indicating self-esteem "becomes an important variable" in interaction with others in the case of masculine perceptions but is much less so for feminine perceptions?
4. What is the meaning of the finding that females with a masculine sex role self-concept perceive a pocket knife to be less masculine than males of any sex role?

Several other such questions may be raised with respect to pocket knife.

Similar questions about implications can be posed in the case of other studied products. For example, what inferences can be drawn from observing that females perceive cuff links to be more masculine and more feminine than do males, or that males perceive hair spray as more masculine and more feminine than do females?

Also it might be asked why relationships involving sex role self-concept, sex, self-esteem, and use vary across the product sex types? Do the notions of sex role self-concept and self-esteem operate in a product specific manner, or do they hold for various products? What do

the results of this study suggest?

A high priority question to resolve prior to undertaking further research, it seems to me, is whether consumers visualize products on a single bi-polar feminine/masculine dimension or on two independent feminine and masculine dimensions as operationalized in this investigation. A study could be designed specifically to answer this question.

The authors state in their conclusions that sex image of products is a bi-polar construct. It was unclear, at least to me, how this conclusion emanated from their data which were collected using separate scales for femininity and masculinity. And if this was the case, why the research analyses examined feminine and masculine product perceptions separately as if they were different dimensions.

The dual versus single dimension question must be answered in order to ascertain whether differences in feminine and masculine product perceptions as well as corresponding relationships with other variables observed in this study were artifacts of dual scales. Such artifacts conceivably could have occurred if dual scales were extremely foreign to the way subjects normally envisioned products.

Another significant research question is whether product perceptions in terms of sex are related (1) to sex role self-concept of the individual or (2) to the sex of persons "most often thought of as using the product." Both relationships were suggested by the current investigators who indicated the latter warranted further investigation.

If the latter relationship exists, then a question arises whether sex role self-concept is a causal variable with respect to product perceptions. Additional questions may be asked about conceptual and/or managerial implications of three-way associations, if they are observed, involving sex role self-concept, perceived sex of typical users, and product perceptions.

Assuming the preceding questions are resolved, research attention might be focused on brand perceptions within product classes. Brand perceptions might differ markedly within a product class that generally has a certain sex image. Cologne, deodorant, and hair spray are examples of such possibilities. Research questions would include whether the variables and relationships relevant across products to sex perceptions also held across brands. The conceptual and managerial implications of the research findings, of course, would need to be identified and defined.

#### Credit Card Usage by Hirschman et al.

One of the study's important contributions was to examine card usage while controlling "external" factors of multiple card possession and retailer acceptance of these alternative cards. Research measuring credit card usage in general terms, aggregated across a variety of purchases and retailers, ignores many situational factors such as retailer acceptance of cards and frequency of buying from each retailer that undoubtedly influence card usage rates.

To ask consumers which card is "usually" used, is to ask them to aggregate their preferences in specific situations, considering retailer acceptance, and to express the summations in terms of which card is used most often. Parenthetically, one might wonder about the ability of consumers to accurately sum across situations and perhaps question the usefulness of aggregate data.

Once research controls are introduced to ensure all re-

spondents possess alternative cards and the retailer will accept each, card usage or preference becomes tied to the prevailing situation. Thus, the comparison of bank and store card user profiles in this study in actuality asks if there are correlates of respondents' card preferences when buying from the particular retailer chosen as the data collection site. Their card preferences in this situation may or may not be similar to those prevailing when purchasing from another retailer.

Given the highly situational nature of the research data, intuitively one might expect demographic factors to have little, if any, discriminating or predictive ability. As the authors indicate, banks and stores are apt to use similar criteria in issuing credit cards. The fact that all respondents were patrons of a particular store probably introduced additional homogeneity into the data because a store's clientele usually has some homogeneity. Consequently, this study's first hypothesis pertaining to demographics assumes the characteristics of a preliminary test prior to turning to a more substantive question concerning correlates of card preferences in this particular situation. Parenthetically, interpretation of the statistically significant variation in income would be hazardous for numerous reasons including its lack of significance when shopping frequency at the store was controlled, as noted by the authors.

There seem to be some unresolved conceptual and managerial questions about the results pertaining to hypotheses 2 and 3. What is the relative significance of mobility factors vis-a-vis perceived importance of card attributes? Would the results have been substantially different if attribute importance ratings had been added to the discriminant function before the mobility factors?

The authors noted that 31% of bank card users were non-residents visiting the city. Length of residency also appeared to be significantly less for bank card users. The amount of absolute difference in residency, however, was not clear.

Many nonresidents and short-term residents might be expected intuitively to have a much lower shopping frequency and lower average monthly credit charges at the store than longer term residents. Would these factors make a difference in the importance attached to certain card attributes such as consolidate billing and utilize additional credit plan? As mentioned earlier, would the mobility factors retain their statistical significance if attribute importance ratings were added to the function first?

In either event, which is more meaningful from the segmentation perspectives of bankers and retailers, to know card attribute importance ratings or to know consumers' location (resident/nonresident) and length of residency?

Another question is what differences would exist between the profiles of bank and store card users if shopping frequency and average monthly charges at the store were controlled.

One possible reason for consumers to have both bank and store cards is to enlarge their line of credit. At least some may think it is easier to obtain two cards with \$500 ceilings, say, than one with a \$1000 limit. Persons living within the area who buy frequently from the store would be in a more advantageous position to exploit this opportunity to extend their credit than nonresidents. Thus, one might ask if perceived importance of the card attribute "utilize additional credit plan" is associated with residence location.

In considering avenues for future research, one might ask what was learned through the current study and what

are significant unresolved conceptual and managerial questions? Several have already been mentioned.

A major research priority, it seems to me, is to sort out the factors that may co-vary with those examined in this study. Such factors would include shopping frequency and average monthly charges. Then it would be important to identify conceptually as well as statistically the most meaningful variables from the standpoint of segmentation.

A longitudinal study might be undertaken to examine if and how card attribute importance ratings and card preferences change over time among consumers who move into the store's area, while again controlling for purchase frequency and monthly charges.

It must be remembered that such analyses are apt to be highly situational, pertaining perhaps to a specific store as in this instance. Thus, further research would be needed encompassing other retailers and purchase situations perhaps before generalizations could be safely drawn about credit card usage preferences.

#### Contingency Data Analysis by Pessemier

This paper makes a valuable contribution in that it summarizes in one place several alternative psychological models of individual judgment and corresponding procedures for identifying and assessing determinant attributes. Its principal concern obviously is with contingency data. Needless to say, it is conceptually based and data free.

Sometimes different results, which at first glance suggest differing interpretations, occur when alternative analytical procedures are applied to the same data base. Consequently, some investigators interested in employing an approach discussed in this paper may wonder if the particular approach selected will influence the results, particularly in the case of contingency data. And if so, they may wish to know how the results might vary and what effects this could have on interpretations. Stated differently, what are the important considerations in choosing an approach? Perhaps if members of the audience are interested in these issues, Pessemier will elaborate on the approaches.

The approaches were discussed from the standpoint of an object rather than that of a judge such as a consumer. For example, the paper tells how an object's attribute complexity and ambiguity might be measured. An implicit assumption in the case of pooled data is that subjects are relatively homogeneous in their judgments, or at least that differences among subjects can be ignored.

Psychology literature suggests, and Pessemier alluded to this earlier in the paper, that persons may vary substantially in their perceptions of an object as well as of similarities/dissimilarities between objects. Some might be said to manifest substantial cognitive complexity utilizing a number of attributes in judging an object. Others appear comparatively simple using only a very few attributes. This raises at least two issues.

One pertains to pooled data. Are results, using the various approaches, sensitive to intersubject variance in judgments? What are the implications?

A second issue deals with differentiating among consumers instead of objects. Pessemier described how some of the approaches could be used to monitor an object over time as it moved through the adoption-diffusion process or product life cycle. Can these same approaches be utilized to segment consumers at a given point in time in terms of attribute judgments; that is to say, in their judgments of a product? If so, how might this be accom-

plished?

#### Closing Remarks

As mentioned earlier, it is hoped that some of these comments and questions might serve to spark research ideas among those hearing and reading the papers. Perhaps certain remarks will raise methodological considerations that in turn will lead to particularly sound research designs and analytical methods in future investigations. Thus, it was not the intent of the points expressed to either applaud or criticize the current work.

In closing, the following general questions might be posed. What were the definitive results, given the concepts, research designs and methodologies employed, and the situational factors prevailing when the data were generated in the case of the empirical studies? Hence, what significant conceptual and managerial questions remain to be answered in the subject areas of the three papers?

William D. Perreault, Jr., University of North Carolina

## Abstract

This is a summary of discussion of three papers which were presented at the "Empirical Methods and Results" session of the 1978 Conference of the Association for Consumer Research. One paper concerns empirical methods for the analysis of contingency data, and the logic behind these methods. Another is an empirical investigation of sex image related product perceptions. The third involves an empirical analysis of alternative predictors of credit card usage behavior. Each of the papers is briefly discussed.

## Introduction

The "Empirical Methods and Findings" session of the 1978 Conference of the Association for Consumer Research focused on three interesting papers. One of these concerns a study of differences and similarities between consumers who prefer two alternative credit card systems: a bank card system and a retail store card system (Hirschman, Srivastava, and Alpert, 1978). A second paper concerns the sex image related product perceptions of a sample of students and how those perceptions vary according to the student's sex, sex role self concept, and several other factors (Golden, Allison, and Clee, 1978). The third concerns different models that have been proposed for analyzing consumer categorical responses to questions concerning the presence of an attribute for a stimulus product (Pessemier, 1978). These papers basically concern very different issues, and thus they will be discussed individually.

## Credit Card Preferences

The study of credit card preferences by Hirschman, Srivastava, and Alpert (1978) appears to have been carefully done, and it is reported in an interesting way. Generally, the analysis suggests that consumers prefer to use a credit system that provides benefits which are most consistent with their life patterns and needs.

The results of the analysis are consistent with expectation. Given that the sample frame is limited to the customers of one relatively "upscale" department store, it is not a surprise that the discriminatory power of the demographic data is limited. In addition, when one considers the specific operationalization of the criterion variable, a discriminating benefit structure between the credit alternatives is logical. Specifically, to be included in the analysis a consumer must actually have had *both* types of credit cards and must also have previously used *both* types for purchases at the store involved. In considering the results of the study, it is important to keep this criterion selection process in mind, as we are dealing with a relatively select group as contrasted with other definitions of card system preference which might yield different group characteristics.

The authors do an interesting job of comparing and contrasting the relative classification efficacy of the various predictors. This is a recurring methodological problem in applied research, yet there are no generally accepted tests for the comparison. The issue underlying such an analysis is the trade off between parsimony and better classification. From this comparative perspective, however, it is possible to directly test the

marginal, multivariate between group discrimination of the more complex model (including both the demographic predictors and benefit variables) as contrasted with the simple (benefits only) model. Given the criterion variable used in this study, it seems that the strength of the benefit predictors is the most relevant theoretical issue. In addition, since the simple model is a nested subset of the total model, the partial F test may be used to determine if discrimination is improved by adding the demographic variables.

In summary, this is an interesting study. Although the generalizability of the specific results to other sample frames may be limited, the study provides a relatively clear empirical example of benefit segmentation. Benefit segmentation research has been approached from many different perspectives, yet frequently the reports in this area are primarily methodological in nature and do not document that the derived segments are "meaningful" in terms of differential market behaviors. Since this study approaches the segments after first identifying the different behaviors, it does provide that evidence in a sample of consumers in which most other potentially confounding factors are, in essence, held constant.

## Product Sexual Images

The study by Golden, Allison, and Clee (1978) documents that there are complex relationships between consumers' perceptions of products, and the consumer's sex, sex role self concept, and self esteem. A particularly interesting aspect of the descriptive analysis in this study concerns the empirical result that for the 125 products initially evaluated none were perceived (rated) as androgenous or neuter by the student respondents. Apparently most products have a relatively distinct sexual identity.

The study is ambitious in its objective of attempting to sort out various factors that may be related to sex-typing product perceptions. As is frequently the case with ambitious undertakings, however, the data are relatively complex. Specifically, from each student respondent repeated measures are taken concerning perceptions of both male and female sexual images of each of five different products; in addition, each of the 307 students is characterized as a member of one of 48 cells in a factorial design. The factorial design is based on the respondent's sex (yes, two levels), sex role self concept (4 levels), self esteem (3 levels), and product usage (2 levels, but the operational definition of these levels is not provided).

Perhaps it is because of this complexity that this reviewer had difficulty understanding the motivation for the research design developed or appreciating the importance of the basic issues which might be addressed by that design. Of course, the authors' relative allocation of emphasis (toward the reporting of data analysis) and away from the theoretical or practical issues relevant to their research) may simply have been a function of space constraints.

At a broad level, the paper stimulates the reader to think about a number of issues relevant to the sexual images of products. For example, the study indirectly raises the question "Can the sexual image of a product be altered by presenting it in alternative use situat-

ions?" A related question is whether or not such perceptual changes would in fact result in altered behavioral usage and consumption patterns. This should be an interesting area for future research. Or, alternatively, does use of a product alter the user's sexual image of that product? Does it alter it in the direction of the user's sexual self concept? Theories of cognitive consistency would predict such results. This would have interesting implications for advertisers of products who were attempting to repioneer their products by opening up new markets for old products by expanding the market segment across lines of gender. This is, of course, the strategy used by many manufacturers of hair spray a number of years ago. If the notion that products have a sexual identity is correct, then this type of positioning effort might be very effective for many companies. For example, if a pocket knife is a masculine product, and presumably the heavy user's are men, then why not a pocket-book knife for women. If one believes that over time advertising has a strong but subtle effect on social values, it might even be that the type of effect could result in some reduction in sex role stereotyping. This is clearly ideal conjecture, and not intended to suggest that these implications are drawn from the authors findings; rather, the intent is to illustrate the types of issues that are related, even if only indirectly to the topical area of this research.

The numerical results and conclusions reported by the authors are also interesting. However, there are methodological concerns relevant to this study which suggests that the authors conclusions and statistical results should be viewed with caution. For example, one must be very cautious in interpreting tests of main effects in the presence of interactions. This concern is especially relevant to nonorthogonal designs, and to designs with small cell size and highly disproportionate cell sizes.

Even ignoring the presence of both low and high order interactions, and the lack of independence of the different measures across the different products, tests of the main effects are likely to be correlated, and the probability levels associated with standard F tests in such a situation are generally not accurate.

These same difficulties may be a partial reason why a clearer pattern of effects does not emerge in the study. With so many cells, there are very few observations in some cells and therefore the response of a single individual may have a substantial impact on results, especially given the multiplicity of responses per subject. More complete descriptive statistics (and cell sizes) would allow the reader to better evaluate whether these concerns are unnecessary. The authors have faced a difficult trade off of reporting a less complete analysis versus a more complex one which is more difficult to interpret.

In summary, this study explores several thought-provoking issues relative to the sex-type image of products. Although the conclusions of the study must be considered with caution due to the complexity of the design and the resulting difficulties in partitioning sources of effects, the study does show that there are generalized differences in the sexual images of products across different respondents, and it stimulates the reader to consider a number of issues relevant to the sexual image of products.

#### Understanding Contingency Data

The paper concerning understanding and analysis of contingency data (Pessemier, 1978) is a very useful review of segments of the literature relevant to the measurement of consumer judgments. The source literature in this area is frequently and unnecessarily complex and

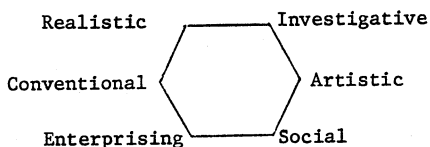
and disjointed. Thus, an effort to concisely integrate the major issues is a substantial challenge, and Pessemier's paper meets that challenge. He not only briefly reviews the major traditional response functions, however, but also provides several insights concerning the response functions. His thoughts concerning consumers' evoked sets, determination of relevant attribute sets, and data "encoding" processes are particularly creative. While other arguments might be advanced to explain some of the same phenomena, the issues Pessemier raises are pertinent not only to those concerned with contingency response data, but more broadly with consumer information processing, measurement theory, and composition models of consumer choice.

Pessemier touches on several interesting topics when he considers perceptual space models as alternatives to the traditional response response models such as the law of comparative judgment or signal detection theory. Although references to appropriate sources are provided, this issue perhaps warrants more detail and emphasis than was provided. A serious limitation of the traditional response models is that they are not suited for individual level analysis, even though they are based on the supposition that there is a distribution of discriminial thresholds. By contrast, recent developments in individual differences multidimensional scaling (and in some forms of monotonic regression) are based on the theoretical basis that the "scale" underlying the individual's response categories may be recaptured, by the scaling process if not by the measurement process itself. These and other issues concerning the relationship between nonmetric multidimensional scaling and the traditional models seems to be very pertinent to potential users and thus to a retrospective synthesis of this sort.

An issue which Pessemier does not address directly, but which also seems pertinent to his synthesis, is the extent to which researchers can quantify the "true" interval scale values which underlie the basic categorical (yes, no, don't know) responses of concern. For example, assume that it is possible to quantify the response categories so that each category (or even each individual within each category) is assigned a numerical value reflecting the relative nature of his or her response. In such a circumstance, the relationship between the response and other relevant behaviors or attitudes of the consumer might be evaluated with more traditional quantitative models. That is, the whole problem of contingency analysis (from the traditional perspective) would be bypassed. In fact, that is the logic motivating the development of a new set of numerical analysis procedures which rely on the principles of alternating least squares optimal scaling. Additional discussion concerning these approaches may be found elsewhere (see, for example, Young, de Leeuw, and Takane, in press; or Perreault and Young, 1978). It seems that these procedures hold much future promise for the type of research situation which Pessemier addresses. More generally, these models for quantifying non-metric data should be relevant to a broad realm of analysis situations, including applications of multivariate linear analysis to mixed measurement level (binary, nominal, ordinal, or interval) data and applications of individual differences metric and nonmetric multidimensional scaling.

In summary, the Pessemier paper simultaneously synthesizes a difficult area and make contributions concerning the relationship of the measurement models to different cognitive areas. It is important that applied researchers consider and understand the behavioral and cognitive (as well as statistical) assumptions underlying quantitative response models; the Pessemier paper, and others like it, are a means for researchers to efficiently gain that perspective.

be social and/or conventional, but they are unlikely to be investigative. Based on his research, Holland has formulated an hexagonal model which shows the degree of consistency between types as being inversely proportional to their distance from one another in the model. The model is depicted below:



Holland (1973, pp. 14-18) describes the six personality types in the following manner:

1. A realistic type prefers activities that entail the explicit, ordered, or systematic manipulation of objects, tools, machines, and animals; and he has an aversion to educational or therapeutic activities.
2. An investigative type prefers activities that entail the observational, symbolic, systematic, and creative investigation of physical, biological and cultural phenomena in order to understand and control such phenomena; and he has an aversion to persuasive, social, and repetitive activities.
3. An artistic type prefers ambiguous, free, unsystematized activities that entail the manipulation of physical, verbal, or human material to create art forms or products; and he has an aversion to explicit, systematic and ordered activities.
4. A social type prefers activities that entail the manipulation of others to inform, train, develop, cure, or enlighten; and he has an aversion to explicit ordered, systematic activities involving materials, tools, or machines.
5. An enterprising type prefers activities that entail the manipulation of others to attain organizational goals or economic gain; and he has an aversion to observational, symbolic, and systematic activities.
6. A conventional type prefers activities that entail the explicit, ordered systematic manipulation of data, such as keeping records, filing materials, and reproducing materials to attain organizational or economic goals; and he has an aversion to ambiguous, free, exploratory, or unsystematized activities.

Holland claims that people will seek congruence between their personalities and their work environments. For example, an individual who is essentially enterprising and conventional will probably choose an enterprising and conventional work environment (which might have some realistic and social characteristics). This person would tend to avoid investigative and artistic work environments (since they would be highly unlikely to have enterprising and conventional characteristics). Thus, an explanation for why many MBA students avoid business doctoral programs may lie in a lack of congruence between the conventional, enterprising personalities of these MBA students and the investigative, artistic nature of the professorial work environment. Likewise, an explanation for why some MBA students choose to enter doctoral programs may lie in the congruence between the social personalities of these students and the social nature of the professorial work environment.

The remainder of this paper is devoted toward reporting on an exploratory study which sought to clarify and test some of the ideas presented in this section. The study was essentially concerned with comparing MBA

students who were not interested in business doctoral programs to both MBA students who were interested and actual doctoral students. It was expected that these groups would differ in ways that the educational and occupational choice literature might predict. That is, they would exhibit different job-market perceptions, family-support situations, intelligence levels, energy levels, personalities, and so on.

#### Methodology

The absence of previous research on why people enter or avoid business doctoral programs dictated that the present study should be exploratory in nature. Given this situation, it did not seem appropriate to incur the expense of acquiring a national probability sample of MBA and doctoral students. It was also deemed infeasible to sample undergraduate students and working people. For achieving the limited objective of clarifying and generating hypotheses, a convenience sample of business graduate students from a representative group of schools was viewed as adequate.

The sample was obtained during the Fall of 1976 by asking professors at thirteen different schools to seek responses from a representative group of MBA and doctoral students (from all majors).<sup>1</sup> Although in some cases questionnaires were distributed in a random fashion (using student mail-boxes), most professors simply handed them out in their graduate classes. The final sample included 236 MBA students and 154 doctoral students.

Students were instructed to complete the study's questionnaire in sequence, and not to go back to a section once it had been completed. The opening question asked the students to indicate the probability of their ever entering a business doctoral program on an 11-point scale (i.e., 0.0 to 1.0).<sup>2</sup> This was followed by three distinct sections. The first section contained a single question which asked the students simply to list their major reasons for wanting and/or not wanting to enter a doctoral program in business or management. This question was included in the hope of obtaining fresh insights about the beliefs, personalities, and other characteristics of the students.

The second section contained a group of 5-point Likert-type scales. The twenty-two scale items were designed to reveal information about the students' (1) job-market perceptions (e.g., "The present supply of people with doctorates in business or management exceeds the demand for these people"), (2) family-support situations (e.g., "My parents would like it if I became a college professor"), (3) personalities (e.g., "Working in the business world requires too much conformity"), and (4) other characteristics (e.g., "I am tired of going to school" and "None of my current friends are interested in entering a doctoral program in business or management").

The final section asked the students questions about their GMAT score, undergraduate grade-point average, marital status, age, sex, and number of years of full-time employment. In addition, questions were asked--based on thoughts obtained from the educational and occupational choice literature--about the size of their undergraduate school, about the social class they grew up in, and about whether they were first-born children.

<sup>1</sup>Responses were obtained from students at the following schools: Berkeley, Case-Western, I.I.T., Indiana, Maryland, Massachusetts, North Carolina, Northwestern, N.Y.U., UCLA, Washington, and Wisconsin.

<sup>2</sup>Doctoral students were instructed to answer 1.0 to this question.



The data from each section were analyzed in separate steps. The responses to the open-ended question in the first section were content analyzed and then cross-tabulated with the responses on the question asking for probability of entering a business doctoral program. The categories for the analysis were developed by the authors and the placing of responses into these categories was done by an MBA student. The responses to the Likert-type items in the second section were first factor analyzed. The four factors which explained most of the variance in the data were then utilized in a multiple discriminant analysis to examine their relationship with the "probability-of-entering" measure. Finally, the association between probability-of-entering and the responses given to each question in the third section were examined using multiple classification analysis, one-way analysis of variance, or correlation analysis, depending on the nature of the data.

## Results

The responses obtained on the portion of the open-ended question which sought reasons against entering a business doctoral program fell within the following three categories:

1. Costs. These responses showed a concern with the time and money it would take to obtain a doctorate.
2. Stress. These responses showed a concern with the amount of work, pressure, and fatigue associated with continuing one's schooling.
3. Career-Goal Incompatibility. These responses indicated that obtaining a doctorate was not useful for achieving their career goals or for getting a good job.

The responses obtained on the portion of the open-ended question which sought reasons for entering a business doctoral program fell within the following four categories:

1. Opportunity to Develop Self. These responses indicated an interest in learning more and in overcoming the "challenge" of getting a doctorate.
2. Opportunity to be Investigative. These responses indicated a desire to become involved with more complex, theoretical research and consulting activities.
3. Opportunity to Teach. These responses indicated a desire to teach.
4. Opportunity for Security, Status, and Income. These responses revealed beliefs that a doctoral degree will lead to job security, increased professional impact, and better incomes and job prospects.

Table 1 reports on the frequency with which different kinds of students gave each type of response. For example, the table shows that 60 out of the 176 negative reasons (34 percent) offered by the least interested MBA students (those who gave a 0.0 to 0.2 probability of entering) were cost-related; and that the respondents who were more interested in doctoral programs gave a larger proportion of cost-related negative reasons. In general, these results suggest that MBA students tend to avoid doctoral programs because (1) becoming a professor is incompatible with their more enterprising and conventional career goals and (2) the costs associated with obtaining a doctorate are perceived as high. The results also suggest that students are attracted to doctoral programs because of a desire to pursue more investigative and social (i.e., teaching) activities. In addition, the results suggest that students who are less interested in doctoral programs are attracted somewhat

by the opportunities these programs present for developing one's intellect or "proving" oneself.

Table 2 presents the results of the factor analysis on the 22 Likert-type items. Using the principal components method with a varimax rotation, four factors were found which explained 77 percent of the variance. To ease interpretation of the factors, consideration was given only to those items which had factor loadings of .4 or greater on one factor and less than .15 on all other factors. The following labels were given to the four factors: (1) Task Competency, (2) Perceived Supply of Ph.D.s, (3) Intolerance for Business Job Activities, and (4) Perceived Demand for MBAs. Items concerned with degree of family support, peer interests, energy level, and other characteristics did not load heavily and/or cleanly on these four factors, suggesting that these variables may not explain much of the differences among business graduate students.

The results of the multiple discriminant analysis are reported in Table 3. This analysis explored the relationship between the four factors (each is measured by summing the item scores for the two highest-loading items associated with it) and the probability-of-entering measure (broken down into the four categories found in Table 1). The bimodal nature of the responses to the probability-of-entering question made it necessary to use a discriminant analysis approach. As the results show, the one significant discriminant function revealed that task competency, intolerance for business activities, and perceived demand for MBAs are positively associated with probability-of-entering, while perceived supply of Ph.D.s is negatively associated. Thus, business doctoral programs seem to look less attractive to students who (1) have doubts about their ability to do professorial-type work, (2) are more tolerant of business job activities, (3) think the Ph.D. job market is poor, and (4) think the MBA job market is poor (reflecting, perhaps, a concern about economic conditions).

The data obtained from the last section of the questionnaire required several different analysis approaches. One-way analysis of variance was used to test whether the four probability-of-entering groups had significantly different GMAT scores, undergraduate grade-point averages, ages, or sizes of undergraduate schools. Doctoral students had significantly higher GMAT scores ( $p = .017$ ) and ages ( $p < .001$ ), but insignificantly lower undergraduate grade-point averages and school sizes. In addition, correlations computed between probability-of-entering and sex, number of years of full-time employment, and childhood social-class produced no significant results, while a marginally significant correlation was found between probability-of-entering and being a first-born child ( $r_{pb} = .09$ ,  $p = .039$ ). Finally, a multiple classification analysis found that students who were married, and especially those who were married with children, reported a significantly higher probability of entering a doctoral program than single students ( $p < .001$ ). This finding can probably be explained by the fact that the doctoral students in the sample were older than the MBA students.

## Discussion

The results obtained from this exploratory study seem to be consistent with Holland's "theory of careers" and with Levine's thoughts about the importance of macro-environmental trends in career choices. The data suggest that an interest in engaging in investigative and social activities, and an intolerance for conventional activities, draws many students toward doctoral programs in business--while opposite interests draw many MBA students toward careers in the business world. The data also suggest that job-market perceptions and perceptions of economic trends draw students toward or away from doctoral

TABLE 1

## REASONS CITED FOR AND AGAINST SEEKING A DOCTORATE IN BUSINESS

	Self-Reported Probability of Entering a Doctoral Program in Business			
	0.0-0.2	0.3-0.5	0.6-0.9	1.0
<b>Reasons Against:</b>				
Costs	60 <sup>a</sup> (34.1%) <sup>b</sup>	38 (46.9%)	10 (50.0%)	39 (55.7%)
Stress	27 (15.3%)	17 (21.0%)	5 (25.0%)	14 (20.0%)
Career-Goal Incompatibility	89 (50.6%)	26 (32.1%)	5 (25.0%)	17 (24.3%)
Total number of negative Reasons Cited by Each Group	<u>176</u>	<u>81</u>	<u>20</u>	<u>70</u>
<b>Reasons For:</b>				
Develop Self	26 (41.9%)	33 (33.3%)	19 (38.8%)	48 (14.9%)
Be Investigative	16 (25.8%)	25 (25.3%)	13 (26.5%)	134 (41.6%)
Teach	16 (25.8%)	24 (24.2%)	12 (24.5%)	108 (33.5%)
Security, Status, and Income	4 (6.5%)	17 (17.2%)	5 (10.2%)	32 (9.9%)
Total Number of Positive Reasons Cited by Each Group	<u>62</u>	<u>99</u>	<u>49</u>	<u>322</u>
Number of People in Each Group	135	71	30	154

<sup>a</sup>Number of cost-related reasons cited by the 135 people in this group.

<sup>b</sup>Percent of the group's 176 negative reasons which were cost-related.

TABLE 2

## FACTOR ANALYSIS OF LIKERT-TYPE ITEMS

Item	Task Competency	Perceived Supply of Ph.D.s	Intolerance for Business Job Activities	Perceived Demand for MBAs
11) I could be a good teacher.	<u>.738</u>	-.034	.009	.006
13) I am capable of writing a doctoral dissertation.	<u>.504</u>	-.095	.016	.091
5) The present supply of people with doctorates in business...exceeds the demand...	-.079	<u>.690</u>	-.032	-.077
6) In 5 years, the supply of people with doctorates in business will exceed demand	-.050	<u>.865</u>	.052	-.052
12) Jobs in industry are boring and repetitious	.015	.006	<u>.659</u>	-.011
22) Working in business requires too much conformity.	.125	-.067	<u>.773</u>	.034
18) The present...demand for MBAs...exceeds the supply.	.074	.038	.028	<u>.641</u>
19) In five years, the...demand for MBAs...will exceed the supply.	.003	.023	.037	<u>.769</u>

TABLE 3

## MULTIPLE DISCRIMINANT ANALYSIS OF PREDICTORS OF PROBABILITY OF ENTERING A BUSINESS DOCTORAL PROGRAM

Factors	Discriminant Weights of Function 1 <sup>a</sup>
Task Competency	.71
Intolerance for Business Job Activities	.44
Perceived Supply of Ph.D.s	-.33
Perceived Demand for MBAs	.13

$$\chi^2 = 167.33; df = 12 (p < .001)$$

$$\omega^2_{\text{mult.}} = .35^b$$

<sup>a</sup>Only one discriminant function was significant.

<sup>b</sup>The omega-squared indicates that 35 percent of the variance in the discriminant function can be accounted for by group membership.

programs. Other factors which seem to have an influence on an individual's decision to seek a doctorate are his skill level (i.e., "task competency," and GMAT score), his ability to withstand the financial hardships of doctoral work, and, perhaps, his age. Variables which do not seem to have a major influence on the decision to enter or avoid a business doctoral program--based on this small, exploratory study--are socioeconomic origin, family support, sex, size of undergraduate school, peer interests, and energy level.

If similar results are found from more conclusive forms of research, then several recruiting strategies would seem in order for schools seeking larger numbers of business doctoral students. One action that could prove helpful would be to publicize the large number of job opportunities available to Ph.D.s in business (as long as these opportunities remain plentiful). It appears that many MBA students see the job market for Ph.D.s as being weaker than it really is.

In addition to providing accurate job-market information, schools seeking doctoral students might find it helpful to communicate the excitement and challenge associated with the research, teaching, and consulting done by business Ph.D.s. Rather than sending out materials and posters dominated by imposing details about degree requirements and faculty publications, it might be preferable for a school to emphasize telling prospective students about the rewarding careers of some of its recent graduates. Efforts should probably be made to target these messages at students who have investigative-artistic-social personalities, rather than at conventional-enterprising types. This might be accomplished by directing messages at students enrolled in certain schools, programs, or courses (e.g., research design, social responsibility).

Another action that would probably help most schools recruit more doctoral students would be to lower the financial and time costs associated with obtaining a doctorate. Besides increasing financial aid packages, some schools might find it feasible to shorten the length of time a prospective student could expect to spend in their program. This might be done without compromising academic standards by setting up rigid time schedules for all students to follow.

## Conclusion

In their efforts to recruit doctoral students, business schools could obtain helpful guidance from additional studies of why people choose to enter or avoid doctoral programs. These studies should probably focus on how personality and macro-environmental trends affect decisions about doctoral programs. In addition to helping recruiting efforts, the insights obtained from these studies could also prove useful for obtaining a more general understanding of educational and occupational choice decisions.<sup>3</sup>

## References

- John L. Holland. Making Vocational Choices: A Theory of Careers. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1973.
- \_\_\_\_\_. The Psychology of Vocational Choice: A Theory of Personality Types and Model Environments. Waltham, Mass.: Blaisdell, 1966.
- \_\_\_\_\_. "A Theory of Vocational Choice," Journal of Counseling Psychology, 6 (Spring 1959), 35-45.
- Adeline Levine. "Educational and Occupational Choice: A Synthesis of Literature from Sociology and Psychology," Journal of Consumer Research, 2 (March 1976), 276-289.
- Gerald Zaltman and Brian Sternthal, eds., Broadening the Concept of Consumer Behavior. Association for Consumer Research, 1975.

<sup>3</sup>The authors would like to express their appreciation for the assistance of Frank Franzak and the twelve professors who distributed questionnaires.

ORGANIZATIONAL BUYING BEHAVIOR: A CONCEPTUAL  
VIEW OF THE BUYING CENTER AS AN INFORMATION  
PROCESSING UNIT

James W. Hanson (student), University of Oregon

Abstract

This paper explores recent developments in the information processing and judgmental decision areas as they may apply to organizational buying behavior. It considers the possibility of expanding information processing concepts to group buying behavior as this may more accurately reflect how decisions are made in organizations.

Introduction

A number of recent articles have discussed the group decision-making process in organizational buying decisions and have made reference to the potential use of information processing research in this area (Webster, 1978 and Wilson, 1978). One of the purposes of this paper is to explore recent developments in the information processing and judgmental decision areas as they may apply to organizational buying behavior. Cognitive theories of information processing are utilized in the fields of psychology, communications and consumer research. The essential focus of such models is on how individuals select, store, evaluate and utilize information for decision making purposes.

Another purpose is to consider the possibility of expanding information processing concepts to group buying behavior. This would follow closely the group dynamics approach to organizational buying. The emphasis is placed on understanding the interaction process among a small group of individuals who may work together, but who probably have different expectations about products and suppliers as well as different decision processes and rules.

Format

Dubin (1976), after outlining the four basic features of a theoretical model (units, laws of interaction, boundaries and system status), indicates there are three other features. These are: propositions - conclusions derived that represent logical and true deductions about the model; empirical indicator - the outcome of converting terms in each proposition; hypothesis - rewording the proposition by substituting the empirical indicator.

This paper will address the proposition and empirical indicator features using a decision process model adapted from Newell and Simon (1972) and Wilson (1977). A series of propositions will be presented and will concern actions of an identifiable group and individuals within that group with respect to:

- |                         |                            |
|-------------------------|----------------------------|
| (a) the buying task     | (f) vendor or supplier     |
| (b) power or influence  | attribute evaluation       |
| (c) roles               | (g) acquisition of sources |
| (d) group size          | of information             |
| (e) conflict resolution | (h) judgmental or decision |
|                         | rules                      |

The empirical indicators will be discussed in terms of possible instruments and methodologies that may be utilized in future research.

The outline for the remaining portion of the paper is:

presentation of the model; description of terminologies; the propositions; suggested instruments and methodologies; and conclusion.

The Model

The decision process model depicted in Figure 1 indicates that both individuals and groups process information and make choices. At the individual level, intervening variables such as perceived risk, buying task, decision style, perceptual bias, and belief structures affect the processing of information and the choice process. At the group level, variables such as buying task, group membership, group roles, power or influence, group interaction, and conflict and conflict resolution affect the processing of information and the choice process.

The term decision process may also be applied to the broader decision making process as depicted by Webster and Wind (1972). Their process includes: problem recognition, information source acquisition; evaluative criteria; supplier identification; and selection process.

The basic elements of the model, then, are an individual and a group in a task environment attempting to choose among sources of information about product-vendor attributes and attempting to choose among a number of product-vendor attribute sets.

Terminologies

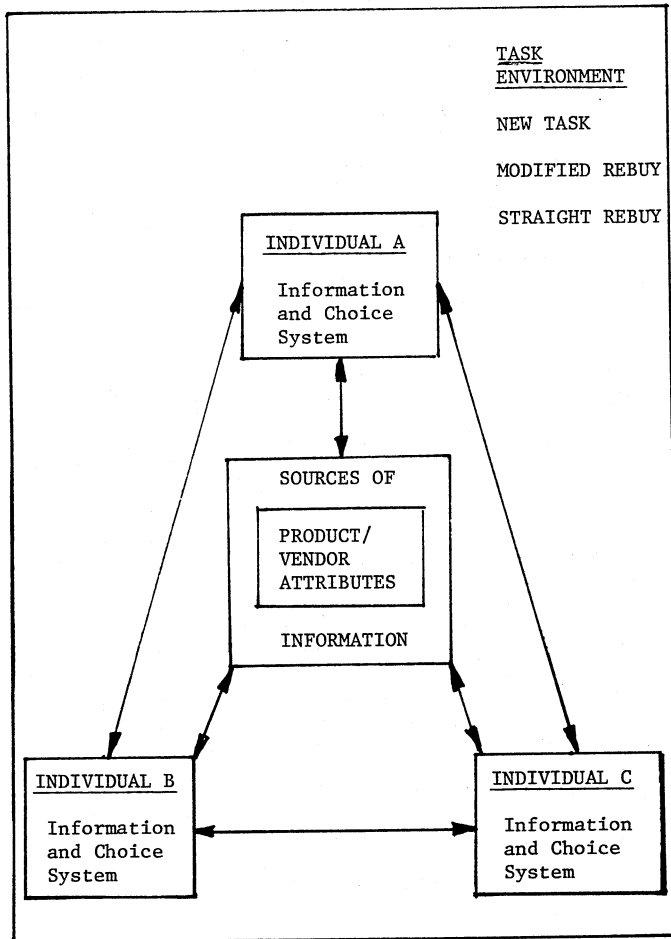
Descriptions, meanings and definitions of various terms, and concepts used in the paper are offered below.

Buying task. There are three types of buying situations: (a) the new task represents the most complex buying situation. The magnitude of the problem-solving challenge, information requirements, and numbers of alternatives will be at a maximum; (b) straight rebuy represents the routine buying situation. The need is not new, the amount of information sought may be small and fewer alternatives may be considered; (c) modified rebuy stands between the two others in terms of newness of the need, the amount of information needed to make a decision and the extent to which new alternatives are considered. (Kotler, 1976)

Buying Center. Those people in the organization who are involved in the buying process may be centrally located, as in a purchasing department, or they may be assembled from various parts of the organization as specific groups or buying committees.

FIGURE 1

DECISION PROCESS MODEL



**Roles.** The buying center consists of five roles: (a) users - those members of the organization who use the purchased products and services, (b) buyers - those with formal responsibility and authority for contracting with suppliers, (c) influencers - those who influence the decision process directly or indirectly by providing information and criteria for evaluating alternative buying actions, (d) deciders - those with authority to choose among alternative buying actions, (e) gatekeepers - those who control the flow of information (and materials) into the buying center. (Webster and Wind, 1972)

**Product-Vendor Attributes.** Various product-vendor attributes lists are available in the literature (Kiser, Rao, and Rao, 1975; Lehman and O'Shaughnessy, 1974; and Wind, Green and Robinson, 1968). These are typically described as relevant choice criteria used by buyers in their choice of vendor or supplier. The lists vary in length from ten to sixty-five.

**Sources of Information.** Different sources of information may be used by various participants in the buying process. Again, lists of information sources have been developed by various researchers (Monoky, Matthews and Wilson, 1975; Ozanne and Churchill, 1968; and Webster, 1970).

**Judgmental or Decision Roles.** These are composition rules used by people in the buying center to process product-vendor related information. Judgmental rules are generally one of two types: (a) noncompensatory-conjunctive, disjunctive and lexicographic and (b) compensatory-unweighted and weighted. Further details of these five rules are given in Park and Lessig (1977).

The Propositions

**Proposition 1:** The Number of People Involved in the Purchase Will Change as the Buying Task Changes.

As the buying task alternates between straight rebuy, modified rebuy and new task, more people, both in numbers and variety, become involved in the decision process. The characteristics of the buying task may account for this as variables such as perceived risk, repetitive versus nonrepetitive purchase, and number and variety of sources of supply vary with the task.

For straight rebuy, decisions may at times be 'automatic' and past experience may be an important factor in supplier selection. The buying center may be composed of only one person who has probably been delegated responsibility for these type purchases.

Modified rebuy tasks entail seeking out more suppliers, seeking more information, and more involvement in actual evaluation of vendors. The product itself may be of a risky nature either from a capital expenditure viewpoint or from its importance in terms of contribution to the production process. More than one person may be involved in the decision process as various points of view from purchasing, finance and production may be required.

New tasks may be 'one shot' deals on a major capital expenditure or of such major significance in terms of impact on the organization, that a great deal of effort and involvement is required for the decision process. The buying center may be composed of representatives from various departments as well as top management in an effort to have diverse views represented.

Few empirical studies have been reported on group size and buying task. Hare (1962) discusses research on size and general tasks such as physical pulling power, concept formation tasks, and completion tasks. With increasing group size, problems may be solved more efficiently, although at some point, depending on the task, the addition of new members brings diminishing returns. Buskirk (1970) and Weigand (1964) quote general statistics on group size and purchase decisions but no breakdown is given with respect to the type of buying task. Hoffman (1965) briefly discusses group size, but only in relation to participation on the part of members. In their summary of a discussion on group size, Ebert and Mitchell (1975) state, "we would expect that routine decisions made at low levels of the organization will be influenced less by group activity than decisions that are more complex and require interpersonal interaction." (p. 173)

There seems to be very little support, both empirically and intuitively, for the notion that buying task and group size are related. Further research is necessary.

**Proposition 1(a):** The Number of Perceived Roles Will Change as the Buying Task Changes.

This proposition follows from the first. As the number of people in the buying center increase or decrease, it seems logical that the numbers of perceived roles will change as well.

As representatives from both purchasing and other areas are included in the buying center, more user, influencer and decider roles are available for interaction. Role sets of individuals may alter depending on the task. These sets are defined by three variables - role expectations, role behavior and role relationships (Webster and Wind, 1972). The interaction process in any given task is partially a result of the role performance of individual members and as such, shifts in the buying task may

bring about shifts in the composition and numbers of role sets.

More specifically, the straight rebuy situation would encompass buyer roles as the decisions are routine. Modified rebuy situations would include buyer, influencer and decider roles as the decision process becomes more complex with the additional search and evaluation stages added. A gate-keeper role may also be present depending on how far the group attempts to go in seeking information and identifying alternative sources of supply. New tasks would encompass the complete range of roles as users, influencers, deciders, buyers and gate-keepers are represented by various organizational interests.

Again there seems to be a lack of empirical research on either identification of specific roles in buying situations or the dynamics of role changes as the buying task changes. Hare does not discuss specific types of roles and only briefly discusses potential role conflicts due to group interaction. Both Hoffman and Hackman and Morris (1975) make no reference to roles or role dynamics at all.

Proposition 2: As Uncertainty of the Buying Task Varies, So Does the Influence of the Appropriate Buying Center.

As the buying task shifts from straight rebuy to new task, the amount of uncertainty faced by the buying center increases. As it has been proposed that the size of the buying center increases with the task, specific persons who constitute the groups associated with the buying center/buying task may become identified as powerful or influential. The primary reason for this association of influence may be the proximity of the buying center to uncertainty. As the uncertainty increases, the individuals who constitute the group designated to handle the uncertainty will be those who are perceived as being influential or powerful.

Research in both the management and marketing areas have explored the relationship between uncertainty and power. Hickson, Hinings, Lee, Schneck and Pennings (1971) put forth the strategic contingencies theory of intraorganizational power and stressed the importance of the ability to cope with uncertainty as a source of power. Salancik, Pfeffer and Kelly (1974) reported on a study of a proposed contingency model of influence in organizational decision making. The method employed had decision contexts which bear close resemblance to the certainty-uncertainty continuum of the straight rebuy-new task buying situations. Their conclusions stated, in part, "as decision making contexts vary, so do the sources of organizational uncertainty and consequently the bases for influence in organizational decision making" (p. 10). Grashof and Thomas (1976) conducted an empirical study on an approach to identifying the varying responsibilities of members of the buying center. Their study also included self and other ratings of influence over stages in the buying process, but not over buying tasks. The only significant result was that individuals tended to inflate their own importance. However, the study does represent an initial attempt at investigating power or influence in an industrial buying context.

Proposition 3: Vendor Attribute Evaluations Will Vary By the Buying Task.

Proposition 3(a): Individuals Within Buying Centers Will Differ With Respect to Vendor Attribute Evaluations.

Proposition 4: Preference For Sources of Information Will Vary By the Buying Task.

Proposition 4(a): Individuals Within Buying Centers Will Differ With Respect to Preference for Sources of Information.

These propositions will be discussed as a group.

As has been shown previously, the buying tasks represent a continuum of uncertainty. High uncertainty may be represented by a lack of information. This lack of information may concern, for example, suppliers, specific attributes of suppliers, specific sources, impact of the decision on the individual and the organization and the consequences of success or failure. Risk is involved and individuals prefer different information and information sources depending on the perceived risk level (Cox, 1967). The propositions reflect, therefore, that sources of information and vendor attribute evaluations will vary by the buying task. Individuals of various backgrounds, from various parts of the organization and with different perceptions will prefer different sources of information and will evaluate vendor attributes differently depending on the buying task.

Propositions 3 and 4: Monoky, et. al, conducted research on information source and vendor attribute preference by industrial buyers as a function of the buying situation. Research instruments were administered to members of the National Association of Purchasing Management. Their results show that preferences for various information sources were systematically related to the nature of the buying task. Also, there were significant differences on the importance of the various attributes across the buying tasks.

Propositions 3(a) and 4(a): Kiser, et. al, concluded from their review of the literature that very little research attention has been paid to the assessment of roles played by relevant non-purchasing executives, i.e., individuals not formally considered part of the purchasing function. These individuals may influence, or may be part of, buying center decisions. Their study set out to find the relative importance of vendor attributes as perceived by non-purchasing executives under special product (new task) and standard product (straight rebuy) buying situations. Their findings indicate there were significant differences between vendor attribute evaluations and buying task.

Both of these studies lend support to the propositions as stated.

Proposition 5: The Buying Center (Group) Vendor Attribute Evaluations Will Vary by the Buying Task and Will Tend to be Similar to Vendor Attribute Evaluations of the Individual Perceived to be Most Influential.

Proposition 6: The Buying Center (Group) Preferences For Sources of Information Will Vary by the Buying Task and Will Tend to be Similar to the Source Preferences of the Individual Perceived to be the Gatekeeper.

These propositions follow closely those stated as propositions 3, 3(a), 4, and 4(a). The common thread through all of these is uncertainty and risk. As the buying task moves along the continuum, the amount of uncertainty and subsequent perceived risk increases. The group interaction processes that may take place as the buying task changes are many and varied. Information will be brought to the group by individuals and the group itself will generate a collective information bank. Throughout this process, information deficiencies will be noted and the group must decide what information to acquire and where to get it.

Individuals identified as being influential or powerful will have an effect on the buying centers actions. The roles involved may be the influencer role or, depending on the type of perceived risk and uncertainty, may be the user or decider roles. These three roles may be represented by one or a few individuals. They would influence the buying centers vendor identification as well as the

vendor attribute evaluations. In terms of sources of information, a fourth role would be influential, i.e., the gatekeeper. Not only does the gatekeeper provide information, he may selectively keep information from the group as well as influence which sources to consider.

Research into the phenomena of 'risky shift' may be of value in discussing these propositions. The findings from recent research have tended to question the hypothesis (individuals confronted with a specific problem will make more risky decisions after participating in a group discussion than they will make alone) but agree that group interaction does produce shifts in judgments. Ebert and Mitchell in their review of the literature on this subject conclude that (a) it is the context of the group process which creates the shift rather than the idea that riskiness is valued or that one feels less responsible and is therefore more risky in a group, and (b) group discussion changes the overall value of the outcomes and the shift reflects a revision of the estimates of the best alternative.

Pettigrew (1972) conducted research on the information filtering process by a gatekeeper during an innovative decision process. The decision process was identified as one involving high uncertainty. The results of the study show conclusively that the gatekeeper does filter information to selected individuals and that this differential access to the flow of information during a decision making process is seen as a source of power.

There has been no evidence of research into group decisions on vendor attributes per se.

**Proposition 7: Judgmental Rules Will Vary by the Buying Task**

One distinguishing characteristic of the stages in a consumer buying process is the consumers familiarity with the product. Howard and Sheth (1969) delineate three stages, (a) extensive problem solving in which predisposition towards and discrimination between brands is low and the consumer actively seeks out information, (b) limited problem solving in which there is a moderate predisposition toward brands, but the consumer does seek out information with which to compare and discriminate between various brands, and (c) routine response behavior in which there is a high predisposition toward one or two brands and very little information seeking.

Park and Lessig (1977) have recently attempted to relate judgmental rules and stages in the decision process. They argue that as the consumer's product and brand familiarity increase, different judgment rules are used for the purpose of identifying those brands which make up the consumer's evoked set and for the purpose of evaluating those brands within the evoked set. The conjunctive model is appropriate for the rejection of alternatives and therefore it is expected to be the one the consumer would use between extensive problem solving and routinized response behavior for the purpose of defining the evoked set composition. The disjunctive model may be more appropriate at the routinized response behavior stage when the consumer may purposefully engage in an exploratory search process to complicate his decision process. This may be seen as overcoming an habitual or boring choice process. Satisfaction of these exploratory needs may come about by accepting an alternative brand with some unique characteristic. The acceptance of a brand is the function of the disjunctive model.

Once the evoked set has been established, the consumer evaluates the alternatives. The unweighted linear compensatory model may be the most appropriate model when the consumer has a low level of familiarity with products and brands. This may be because he has established neither a related set of evaluative criteria to be used in product evaluations, nor a specific set of attribute

weights. The weighted linear compensatory model may be the most appropriate at moderate and high levels of familiarity. The consumer is expected to have established both a reduced set of evaluative criteria and a corresponding weighting function. From this, it is possible to match the stage and the rule. The consumer will use the unweighted linear compensatory model in the extensive problem solving stage and the weighted linear compensatory model as he moves through the limited problem solving and routine response behavior stages.

The lexicographic model is assumed to be utilized in the routine response behavior stage. The consumer has well defined evaluative criteria and high familiarity and so can select a brand which is superior on a most important attribute.

The concept of stages in the consumer decision process is similar to the organizational buying task continuum. The extended problem solving stage can be matched with the new task; the limited problem solving stage matched with modified rebuy; and routine response behavior stage matched with straight rebuy. It can now be postulated which judgmental rules may be used at each part of the buying task continuum.

New task - the uncertainty element of this task may be made up of a lack of information about suppliers attributes, a lack of knowledge about the product itself and which of its attributes may be critical for the organization. These characteristics would imply that the unweighted linear compensatory model may be used by the buying center in its evaluation and choice processes. Individuals may show varying types of judgmental rules in this type of task, but the group interaction process would necessitate at least an accommodation of viewpoints from individuals who may represent areas of the organization other than the purchasing function.

Modified rebuy - a lesser amount of uncertainty may exist as the buying center has had experience with previous purchases of this nature. Familiarity is increased and there may be a set of evaluative criteria established with appropriate weights. Individuals within the buying center for this type of task again may exhibit use of varying judgmental rules but the characteristics of the situation imply that a conjunctive model and a weighted linear compensatory model may be in use. The conjunctive model may be used to reduce the numbers of suppliers the buying center considers in making a choice and the weighted linear compensatory model used to choose between suppliers who remain in the 'evoked' supplier set.

Straight rebuy - the automatic repurchase characteristic implies less uncertainty and a good deal of information and past experience with both the product and selected suppliers. Again a weighted linear compensatory model is most likely to be used in this buying task by the individual or relatively few individuals involved. Both the evaluative criteria and the list of suppliers has been limited to a select and important few. Two other models may appear in this task. The lexicographic model may be used if a critical attribute has been identified. The select set of suppliers, even though small, may then be evaluated against this attribute. Also, the disjunctive model may at times be used if the individual(s) responsible for purchasing intentionally complicates the buying process by considering a new supplier or new product.

**Proposition 8: Types of Conflict and Conflict Resolution Approaches Vary by the Buying Task**

Conflict, according to March and Simon (1958), is present when there is a need to decide jointly among a group of people who have at the same time, different goals and perceptions. The buying task continuum indi-

cates that various individuals comprise buying centers and therefore it is highly likely that conflict is a common consequence of decision making within these centers. Sheth (1973) points out that there are both rational and irrational methods of conflict resolution which may be applied when specific types of conflict arise:

- (a) if the inter-party conflict is largely due to disagreements on expectations about the suppliers or their products, the conflict will be resolved in a problem-solving manner,
- (b) if the conflict among the parties is due to disagreement on specific criteria with which to evaluate suppliers - although there is an agreement on the buying goals or objectives at a more fundamental level - it is likely to be resolved by persuasion,
- (c) if conflict arises due to fundamental differences in buying goals or objectives among the various parties, the conflict will be resolved by the process of bargaining,
- (d) if the disagreement is not simply with respect to style of decision making, the conflict boarders on the mutual dislike of personalities among the individual decision makers. The resolution of this type of conflict is usually by political tactics.

Conflict may be present in the various buying tasks. The conflict of type (a) may be characteristic of new task and modified rebuy. The problem solving approach may be indicative of the buying centers use of compensatory type judgmental rules. Conflicts of type (b) may be characteristic of straight rebuy and the fact that a lexicographic judgmental rule may be used by one or more individuals. If the important attribute differs by individual, then conflict arises. Resolution by persuasion may just involve consensus on which of the two important attributes should be used in the lexicographic model. The conflicts of type (c) and (d) result in irrational methods of resolution and may reflect behaviors of powerful or influential individuals. These may arise in both the new task and modified rebuy as the roles of user, decider, influencer and gatekeeper interact.

Cyert and March (1963) offer the concept of 'quasi resolution of conflict' as a rational means of conflict resolution. The mechanisms suggested to reduce conflict are local rationality, acceptable level decision rules and sequential attention to goals. Strauss (1962) studied tactics used by purchasing agents to influence their relationships with other departments. These tactics may be thought of as representing the irrational methods of conflict resolution.

#### Suggested Instruments and Methodologies

Pretests to determine characteristics of buying tasks should first be conducted. Then questionnaires or personal interviews may be used to identify individuals involved in each task. Tabulations and simple summary statistics would indicate the composition of the buying center. Individuals may then be asked to identify particular people as to the role they play in various buying tasks. Cross tabulations and appropriate chi-square calculations would indicate relationships between self reporting and reporting of others for particular roles.

With respect to perceived roles and the buying task, participants may be asked to rank all participants in order of their importance in making decisions. The summed ranks could be used as a measure of influence. Ken-

dalls measure of concordance may be used to assess the agreement among rankings (Salencik, Pfeffer and Kelly, 1974). Another approach again uses rank order of specific individuals and may be applied across buying tasks. The difference lies in use of Thurston's Law of Comparative Judgment (Case V) to develop interval scaled data from rank order data. An approach might be to locate individuals position of influence on interval scales according to their influence in each of the three buying tasks (Grashof and Thomas, 1976).

Pretests to determine sources of information normally used by particular organizations and attributes used in evaluating suppliers should be carried out before any study. Attitudinal type models such as those developed by Fishbein and Rosenberg may be used to compute both summary indices of preference toward a supplier and composite indices of the importance of attributes and perceived ability of a supplier to provide specific valued attributes.

Preference rating scales for various sources of information in a particular situation and importance rating scales for attributes in particular situations may be administered to individuals and then to groups. The group approach might employ verbal instructions as to what the group should accomplish. Interaction processes as well as the final outcome with respect to sources of information and evaluative criteria may be recorded.

Roles of individuals recorded earlier may be then compared to both individual and group decisions. Appropriate non-parametric tests may be applied for analysis.

Various methodologies have been used to identify individuals usage of judgmental rules. The 'think aloud' or verbal protocol method has been used in consumer research to identify what judgment rules are being used in the process of information seeking (Payne, 1976). Individuals as well as groups may participate in this type of exercise. Recordings of verbal responses either by coders or by video tape could be content analyzed to determine the categories of judgmental rules by task.

Wright (1975) suggests another methodology and that is to fix the judgmental rule and ask individuals or groups to indicate how they would apply the rule to various problem situations. Measures of precision and accuracy may be obtained as well as attitudinal and preference data with respect to different types of judgmental rules.

Group interaction processes would have to be studied in order to determine the types of conflict under various buying task situations. Methods of resolution would also be recorded, again either by coders or by first video taping group sessions and content analyzing them after the fact.

In general, the types of comparisons required under these prepositions may necessitate use of laboratory type experiments but in an organizational setting. Groups of individuals involved in purchasing may be asked to participate and appropriate experimental designs would have to be developed to accommodate individual and group comparisons. Internal validity would be heightened by this approach but external validity or generalization may be limited. Field experimentation would enhance external validity but due to the requirements of measurement, may not be appropriate. Problems with organization participation, measurement and response limitations would limit this approach to individual firms. Longitudinal studies in a single organization would provide relevant duties for testing these propositions. Again participation would be a problem and generalization to other firms and other buying centers would not be viable.



## Conclusion

A series of propositions concerning actions of an identifiable group and individuals within that group have been presented in an organization buying context. Suggestions were offered with respect to methodologies, instruments and empirical measures that might be used in testing these propositions.

Some research has been conducted into a few areas indicated by the propositions but the overall aspect of comparisons between group and individual decision processes has not been looked into by researchers. Specifically, research into group approaches to vendor and source of information selection under varying buying task situations is lacking. Very little research has been done on judgmental rules used either by individuals or groups in an organizational buying behavior content. The dynamics of role set formulation as the buying task varies also has not been studied. By identifying gaps in current and past research, there is both an implicit and explicit call for research in the areas discussed.

## References

- Richard H. Buskirk, Principles of Marketing, New York: Holt, Rinehart and Winston, 1970.
- Donald F. Cox, ed., Risk Taking and Information Handling in Consumer Behavior, Graduate School of Business Administration, Harvard University, 1967.
- Richard M. Cyert and James G. March, A Behavioral Theory of the Firm, Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1963.
- Robert Dubin, "Theory Building in Applied Areas," in M. Dunnette (ed.), Handbook of Industrial and Organizational Psychology, Chicago: Rand McNally, 1976.
- Ronald J. Ebert and Terrance R. Mitchell, Organizational Decision Processes, New York: Crane, Russak and Company, Inc., 1975.
- John F. Grashof and Gloria P. Thomas, "Industrial Buying Center Responsibilities: Self Versus Other Member Evaluations of Importance," in Kenneth L. Bernhart, (ed.), Marketing: 1776-1976 and Beyond, 1976 Proceedings Series #39, (Memphis: American Marketing Association, 1976), 344-347.
- J. Hackman and C. Morris, "Group Tasks, Group Interaction Process, and Group Performance: A Review and Proposed Integration," in L. Berkowitz (ed.), Advances in Experimental Social Psychology, New York: Academic Press, 1975 (Volume 8).
- A. Hare, Handbook of Small Group Research, New York: Free Press, 1962.
- D. Hickson, C. Hinnings, C. Lee, R. Schneck and J. Pennings, "A Strategic Contingencies Theory of Intraorganizational Power," Administrative Science Quarterly, 16, 1974, 216-229.
- L. Hoffman, "Group Problem Solving," in L. Berkowitz (ed.), Advances in Experimental Social Psychology, New York: Academic Press, 1965 (Volume 2).
- John A. Howard and Jagdish N. Sheth, The Theory of Buyer Behavior, John Wiley and Sons, 1969.
- C. Kiser, C.P. Rao and S.R.G. Rao, "Vendor Attribute Evaluations of Buying Center Members Other Than Purchasing Executives," Industrial Marketing Management, 4, 1975, 45-54.
- Philip Kotler, Marketing Management: Analysis, Planning and Control, 3rd Edition, Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1976.
- Donald Lehman and John O'Shaughnessy, "Difference in Attribute Importance for Different Industrial Products," Journal of Marketing, 38, (April, 1974), 36-42.
- James G. March and H.H. Simon, Organizations, New York: John Wiley and Sons, 1958.
- John F. Monoky, Jr., H. Lee Matthews and David T. Wilson, "Information Source Preference by Industrial Buyers as a Function of the Buying Situation," Working Paper #27, Working Series in Marketing Research, College of Business Administration, The Pennsylvania State University, 1975.
- A. Newell and H.A. Simon, Human Problem Solving, Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1972.
- Urban B. Ozanne and Gilbert A. Churchill, "Adoption Research: Information Sources in the Industrial Purchasing Situation," in Robert L. King (ed.), Marketing and the New Science of Planning (American Marketing Association), 1968.
- C. Park and V. Parker Lessig, "Judgmental Rules and Stages of the Familiarity Curve: Promotional Implications," Journal of Advertising, 6, 1977, 10-16.
- John W. Payne, "Heuristic Search Processes in Decision Making," in B.B. Anderson (ed.), Advances in Consumer Research, Vol. III, (Cincinnati: Association for Consumer Research, 1976), 321-327.
- A. Pettigrew, "Information Control as a Power Resource," Sociology, 6, 1972, 3-11.
- A. Salancik, J. Pfeffer and J.D. Kelly, "A Contingency Model of Influence in Organizational Decision Making." Paper presented at 34th Annual Meeting of the Academy of Management, Seattle, Washington, August, 1974.
- Jagdish N. Sheth, "A Model of Industrial Buyer Behavior," Journal of Marketing, 37, (October, 1973), 50-56.
- George Strauss, "Tactics of Lateral Relationship," Administrative Science Quarterly, 7, 1962, 161-186.
- Frederick C. Webster, Jr., and Yoram Wind, "A General Model for Understanding Organizational Buying Behavior," Journal of Marketing, 36, (April, 1972), 12-19.
- Frederick C. Webster, Jr., "Management Science in Industrial Marketing," Journal of Marketing, 42, (January, 1978), 21-27.
- Robert E. Weigand, "Why Studying the Purchasing Agent is Not Enough," Journal of Marketing, 32, (January, 1968), 41-45.
- David T. Wilson, "Dyadic Interactions," in Arch G. Woodside, Jagdish N. Sheth, and Peter D. Bennett (eds.), Consumer and Industrial Buying Behavior, New York: North Holland, 1977, 355-365.
- David T. Wilson, "Research Approaches to Multi-Participant Decision Process," Keith H. Hunt, (ed.), Advances in Consumer Research, Volume V, (Ann Arbor: Association for Consumer Research, 1978), 639-642.
- Yoram Wind, Paul C. Green and Patrick J. Robinson "The Determinants of Vendor Selection: The Evaluation Function Approach," Journal of Purchasing, 1968, 29-41.
- Peter Wright, "Consumer Choice Strategies: Simplifying vs. Optimizing," Journal of Marketing Research, XII, (February, 1975), 60-67.

# INNOVATIVE HEALTH CARE ALTERNATIVES:

## A MODEL OF THE CONSUMER BEHAVIOR PROCESS

Stephen J. Miller, Oklahoma State University  
William G. Zikmund, Oklahoma State University

### Abstract

The choice among health care alternatives by the individual is a consumer decision process. The authors integrate prior health care research into a schematic model of the decision process. Then they empirically examine the model constructs as explanation of intentions to patronize a nurse-practitioner.

### Introduction

Our nation is increasingly becoming aware of the need for improved health behavior and health care. To this end, solutions of a "production" orientation have emerged with emphasis on financial support for more doctors, nurses, hospitals, and insurance plans. Yet there exist continuing problems of inadequate health care, particularly among disadvantaged population segments such as the poor, ethnic minorities, the lower educated, and the elderly.

Long range solutions will likely require innovative delivery methods. The most publicized of these are the nurse-practitioner and physician-assistant programs. Other delivery components might be the mobile medical units of a "circuit rider" nature, day clinics and emergency helicopter service. While the medical profession may view these activities as accepted delivery system components, to the typical user/consumer, the physician is the primary entrance mechanism to the health care system. Alternatives often are perceived as innovations fraught with uncertainties, risks, and sundry misgivings. The development of such delivery system alternatives will be no means assure their acceptance by the public.

Only recently have social scientists come to realize the need for a "health care consumer" orientation (e.g., Aday and Andersen 1975; Andreopoulos 1974; Freeman, Levine, and Reeder 1972) as integral to improving the health of the public. Evidence has begun to accumulate indicating that social-psychological factors influence the range of decision involved in preventive, diagnostic, and therapeutic care. Therefore, the behavioral forces at work on the demand side of the health care equation must be given serious attention.

### RESEARCH OBJECTIVES

Rural America represents a prime focus of attention for disadvantaged consumers regarding access to health care (Andreopoulos 1975; Luft, Hershey and Morrell 1976). Disproportionate numbers of poor, elderly, and low educated individuals comprise the population. Likewise, distances of 30-40 miles to health care personnel and facilities, antiquated facilities, and "overworked" elderly physicians are common occurrences. Yet, the rural consumers may very well reject the innovative health care delivery system necessary to assure their access to primary care. Attention must turn to the psychological factors that influence their acceptance or rejection of such offerings.

This research reports on a study of the choice regarding the decision to patronize a nurse-practitioner lo-

cated in a rural community for diagnostic or therapeutic care versus the alternatives of physician (or perhaps direct avoidance of medical aid). The research objectives that guided the study were:

1. How receptive are the individuals to a nurse-practitioner in the community?
2. What socio-economic and psychological factors explain variability in willingness to patronize a nurse-practitioner?
3. How do the factors that enter the patronage intention decision differ by sex?

More specifically, the research examines a number of constructs felt to be relevant components of the decision. The model uses constructs from the field of medical "sociology" and theoretical constructs in consumer behavior. (The work of Suchman 1966; Rosenstock 1966; Zaltman and Vertinsky 1971; and Howard and Sheth 1969 are the prime base for the model).

Although health care decision making can be viewed in the traditional consumer behavior literature, the operational concepts in brand choice situations have different meanings in their usage across specific situations. Thus, generalized expectancies should not necessarily be formulated at the highest level of abstraction (e.g., personality). Instead, domain-specific expectancies such as "medical fatalism" can be more fruitfully intergrated into the health care model. Exploratory development of measures for such domain-specific constructs have been utilized in the research.

The section to follow will examine the literature germane to this research and the relevant constructs to be examined. Then the empirical measurements will be explained. Finally, the results of the study are presented.

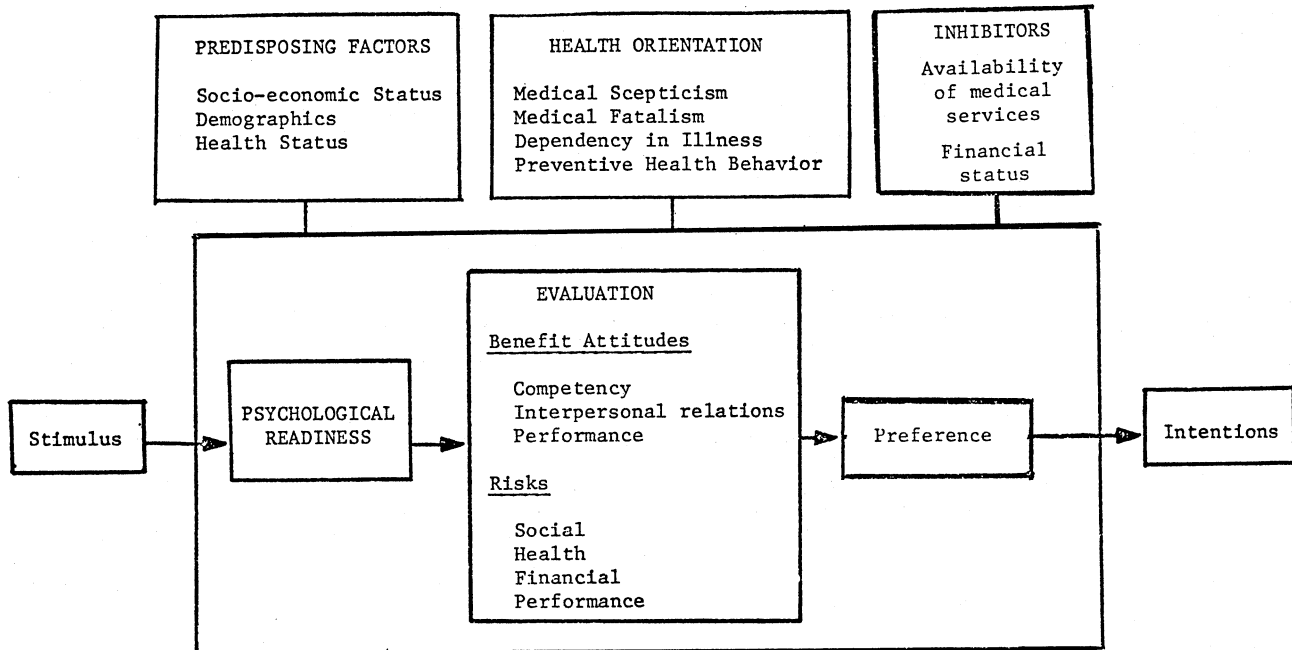
### A Model of Health Care Utilization

A great deal of research has been conducted relative to utilization of health services. Selected reviews of the more important research are presented by Aday (1972) and by Andersen and Bartkus (1973). Based on this past research, the health care model for the research has been developed. Figure A schematically portrays the variables in this model. The model focuses on a decision regarding a health care innovation. Thus, it is oriented to the fundamental decision whether or not to utilize the innovation. The proximal variables are hypothetical constructs within the domain of the specific decision process. The distal variables are general characteristics and predispositions that provide the context within which health care decisions are made. Of course, there is a specificity-generality continuum along which all the variables lie. Proximal variables will be considered first in the model discussion.

The decision process in the medical care choice will be activated by receipt of stimuli. A decision maker might receive a variety of stimuli that activate decision making. Any given stimulus might be a physiological state such as pain, a social comment such as "you look pale today" or a myriad of other stimuli.

Receptivity to the stimuli is a function of the psycho-

FIGURE A



logical readiness on the part of the decision maker. Given readiness, the evaluation of alternative responses will follow. The evaluation will include a combination of perceived benefits (attitudes) and risks. If uncertainties exist, search for information will occur. Likewise, a variety of inhibitors (barriers) exist that will mediate the attitude role as it influences the behavioral intention. The key components of the evaluation process are discussed below.

Symptom Sensitivity

The immediacy and strength of response to the stimulus will relate to psychological readiness. Readiness has been defined as a two-component dimension by Rosenstock (1966) that reflects sensitivity to health-based stimuli as a mediating factor. It is a composite of perceived susceptibility to the given condition and seriousness of the condition. The susceptibility component is the subjective probability of illness contraction as perceived by the individual. The seriousness dimension is the perceived consequence, either medical, economic, or social. For example, an individual may view cancer as a low probability disease with a very serious consequence. In contrast, a headache may appear moderately likely with little concern for its seriousness.

The psychological readiness construct has less specificity than others considered as proximal variables. However, it is discussed at this point because of its expected initial role in the decision process.

Health Care Attitudes

A decision to act given receipt of a stimulus primarily focuses on the perceived benefits of taking such action (c.f., Rosenstock, 1966). Such perceived benefits have been particularly valuable in the study of preventive health behavior. For example, a relationship between perceived benefits and behavior has been noted for chest X-rays (Hochbaum, 1958), dental check-ups (Kegeles, 1963) and inoculations (Leventhal et al., 1960).

The benefit attitudes are by nature domain-specific. In this research instance, they will deal with beliefs regarding the effectiveness of utilizing the health service innovation for treatment of a perceived illness. Benefits in the case of nurse-practitioners might include items such as reduction in waiting time, lower travel or fee costs, greater empathy, and relative medical competency.

Using a factor analysis, 12 specific beliefs toward nurse-practitioners were reduced to three main dimensions or benefit factors. These are competency of the nurse-practitioner to perform the task, interpersonal relations with the patient, and comparative performance of the task relative to the more traditional physician (Zikmund and Miller, 1977).

Perceived Risk

In conjunction with perceived benefits, the decision maker will likely consider the possible negative dimensions of treatment by a nurse-practitioner, treatment risk. This construct has strong empirical validation in the consumer behavior literature. First introduced by Bauer (1960), he remarked that "consumer behavior involves risk in the sense that any action of a consumer will predict consequences which he cannot anticipate with anything approximating certainty, and some of which, at least, are likely to be unpleasant." Thus, a consumer's perceived risk is composed of (1) the subjective uncertainty as to the outcome of actions and (2) the possible consequences related to any given action. Subsequent investigations have viewed perceived risk as multi-dimensional phenomena (e.g., Zikmund and Scott, 1973).

Four major types of risk will most likely apply to nurse-practitioner consideration by the individual. It is believed that there will be a financial risk of utilizing a nurse practitioner. That is, it may not be worth the money to visit a nurse-practitioner or it may be a waste of money. There may be a social risk in the sense that a person's friend or significant others' appraisal of a

nurse practitioner may be perceived to be negative (Franklin and McLemore, 1970). There is further a performance risk in the sense that a nurse practitioner may not perform as skillfully as a medical doctor and perhaps may cause pain to the individual. Finally, there is a health risk in the sense that the person's condition may remain the same or even worsen. Other forms of risk such as convenience risk may be important in the sense that there may be a time loss by visiting a nurse practitioner. However, these dimensions of risk are believed to be of less importance.

#### Inhibiting Factors

A wide variety of inhibiting factors might be considered. These could be included as either proximal or distal to the decision process depending on how situation-specific the variables are defined.

The ready availability of services influences utilization of health services (Bice and White, 1969). Of particular relevance in a model of nurse practitioner selection is the factor of distance, since the nurse-practitioner study focuses on rural areas. The distance component would include both physician and hospital dimensions of availability.

The method by which one pays for medical care is an inhibitor of action. Income is discussed later as a distal variable. However, directly health-linked payment methods can be viewed as inhibitors. For example, Klarman (1963) and Andersen and Sheatsley (1959) both found lower usage rates among prepaid group practice plans than among fee for service plans.

#### General Influences

The distal variables are those of a general nature that influence the decision process as structural factors. Their influence is expected to be less direct than that exerted by the proximal variables. The major categories of distal variables are discussed below.

#### Socio-demographic Influences

A number of variables related to socio-economic status have been found relevant to health care decision making. Suchman in a series of articles (1966a; 1966b) investigated ethnic and social factors in medical orientation. He found that social group structure, i.e., cosmopolitan versus parochial groups differed in medical orientation. He also found that certain demographic factors such as age, sex, and social class influenced medical orientation, knowledge about disease, skepticism of medical care, and dependency in illness. Further, demographic factors, medical orientation, and health status influence the source of medical care (1966b). Subsequent research supports the influence of socio-economic status variables (c.f., Greersen et al., 1975).

In a behavioral model of health services developed by Andersen (1968), predisposing variables that established a health care orientation and need were identified. These included socio-demographic and social-psychological correlates. Finally, age is an important predictor of health utilization, and the relationship between physician utilization and age tends to be U-shaped, i.e., old people and very young people tend to use more services. Education increases the probability of overall physician usage, particularly with respect to preventive health care behavior. Marital status, family size, race and ethnicity have also shown some influences on physician utilization. Of particular significance in understanding health care decision processes is the likely structural differences in the decision process by sex. It has been noted that men and women encounter the need for health services at different stages in their lives

(Booth and Babchuck, 1972). Also, the woman usually plays the health care decision maker for the family as a whole, particularly for herself and the children.

Under the label of enabling factors as posed by Andersen (1968), economic status has shown significant relationships to physician utilization. For example, the higher the income, the greater the utilization of physician services. Price of medical services and methods of financing are related to income and appear to be important factors.

#### Health Orientation

The individual possesses a number of general opinions concerning health and health care matters. These are drawn from past experience and general social influences. Suchman (1966a) has found medical scepticism to play an influence on the health care decision. The scepticism construct reflects an uncertainty as to beneficial returns from the medical profession and health care behavior. Suchman also found a psychological dependency in illness construct to reflect a health orientation. In past studies, these have not been extremely valuable in predicting the use of health services (Aday, 1972, p. 21). However, they are included in the current analysis since the consumer decision being investigated is an innovative choice.

A health orientation variable that appears relevant to the nurse practitioner decision is medical fatalism. This construct is based on Zaltman and Vertinsky's (1971) fatalism and is in line with Rotter's (1966) notion of internal-external control of reinforcement. It is believed that certain individuals will hold that being in good health depends more on being lucky, chance, etc., than on any activities within the control of the individual. The more fatalistic an individual is, the less likely he is to see any benefits resulting from his or an agency's efforts to improve his health status.

The health history should have a variety of influences on the health care decision. This is partially felt through the effect of health history on the health orientation as learned responses. However, direct influences may well occur (Bice and White, 1969; Wirick, 1966; Andersen and Bartkus, 1973). One would expect that a history of frequent physical exams, a family doctor, and other links to an established delivery system would negate receptivity to alternative delivery systems. On the other hand, a frequency of past illnesses would likely stimulate general interest in an expanded system.

#### Research Method

The research data for examining the model construct of FIGURE A were obtained through in-home personal interviews with a cross-section of adults residing in small Oklahoma communities. These rural areas include concentrations of poor, elderly, and lower educated citizens. Thus, an environment exists that is conducive to alternative health care delivery systems.

In conjunction with representatives from the Oklahoma Health Planning Commission staff, a judgment sample of 10 communities in Oklahoma were selected for study. Criteria in the selection were: communities of approximately 2500 residents or less; diverse health care delivery system capacities over communities (ranging from no physician-hospital to two physicians-hospital); geographic and socio-economic diversity over communities. Visual surveys of housing types and conditions were conducted within each community then households were selected for inclusion in the study on a purposive basis to yield socio-economic diversity. In total, 220 adult household members were interviewed. Women were sampled

more heavily since they have traditionally played the greater role in household health care decisions. From the original sample, 205 usable questionnaires were obtained (148 women, 57 men).

The data collection was by a personal interview with a variety of question formats, both open and closed ended and required approximately 45 minutes to administer. Pretesting of the questionnaire for clarity and ease of administration was conducted in a non-surveyed community.

The measurements for model constructs are given in FIGURE B. Most are standard measures (e.g., Medical Scepticism, Psychological Readiness, Dependency in Illness, Health History and Demographics). Others specific to nurse-practitioners (e.g., benefits and risks) and the Medical Fatalism scale have been constructed for this research.

Prior to the questions linking directly to nurse practitioners, the respondents were read a description that explained the training of such medical personnel, their general treatment activities and a plan for the nurse practitioner to maintain a primary care office in the community while working closely with a physician in a nearby town for referral and advice.

### Results

The respondents were queried regarding familiarity with the nurse practitioner concept along with two other innovative delivery methods to disguise the study's purpose. Familiarity and general orientation toward the medical innovation are shown in TABLE 1. There was little initial familiarity. Forty-five percent of the respondents had never heard of a nurse practitioner while another 31 percent possessed only a basic recognition. Thus, only 24 percent had any explicit cognitions linked to the concept.

Willingness to visit a nurse practitioner was considerably more favorable than the familiarity status. Note that 73 percent express some degree of willingness while only 17 percent reject the concept.

The "willingness to visit" is general. As such, it does not measure the intensity of patronage intention over a variety of situations. Patronage intent was measured based on anticipated response to an array of fifteen symptoms from minor to major significance (e.g., fall, chest pains, high blood pressure). The intention results are presented in TABLE 2 for both men and women. These indicate the average number of medical symptoms for which the individual expressed an intention to visit a nurse practitioner for treatment should one be available.

TABLE 1

#### RECEPTIVITY TO NURSE PRACTITIONER

Knowledge Response	%	Willingness to Visit Response	%
Never heard of it	45	Definitely would not	4
Heard of it	31	Probably would not	13
Know just a little	14	Uncertain	10
Know a fair amount	6	Probably would	56
Know very much	4	Definitely would	17

For the fifteen symptoms included in the query, the average number for which a respondent would patronize a nurse-practitioner was slightly more than 50 percent for women and less than 50 percent for men. The differences

in patronage intention for women and men were statistically significant at the .001 level with the results indicating less receptivity to nurse practitioners by men than by women. This should likely be interpreted as a relative perspective since women tend to be more habitual health care users than men due to pregnancy, pediatric care for children, menopause, and other health care needs throughout their lives.

FIGURE B

#### MODEL VARIABLES

Variable	Description
Patronage intention	Summation of <u>intentions</u> (yes=1) to visit an available nurse-practitioner; 15 different medical symptoms (e.g., cut, fall, chest pains, ear infection).
Psychological readiness	Summation of perceived likelihood x seriousness for five health conditions with measures on 5 point scales.
Benefit attitudes	Summations of beliefs with regard to nurse-practitioner <u>competency</u> , <u>inter-personal relations</u> with patients, and <u>treatment performance</u> skills. Each attitude cluster represents a set of four beliefs measured on 5 point scales.
Perceived risk	Summation of perceived importance x result likelihood for <u>social</u> , <u>financial</u> , <u>health</u> and <u>performance risks</u> measured on 5 point scales.
Inhibitory factors	<u>Physician</u> and <u>hospital availability</u> in community (number), <u>distance</u> to closest <u>physician</u> and <u>hospital</u> (miles) and participation in <u>health insurance plans</u> (yes=1).
Socio-economic factors	Respondent <u>age</u> , <u>education level</u> , number of <u>individuals under 18</u> in household and annual <u>household income</u> .
Health orientation	Summation of beliefs with regard to <u>medical scepticism</u> (trust of medical profession with high being sceptic, 3 beliefs), <u>dependency in illness</u> (need for personal concern with high being dependent, 1 belief), and <u>medical fatalism</u> (felt personal control of health status with high being internal control, 5 beliefs). Measured on a 5 point scale.
Health history	Maintenance of <u>family doctor</u> , <u>recent illness</u> , <u>physical exam</u> , or <u>hospitalization</u> (yes=1).

TABLE 2

#### PATRONAGE INTENTION

Women		Men	
Mean	Std. Dev.	Mean	Std. Dev.
8.04*	3.94	6.93*	4.37

\*Significant difference at .0001

## Structural Correlates

Each variable in the model was correlated with patronage intention. These correlations are given in TABLE 3. Of the forty-eight correlations, nineteen are statistically significant at the .1 level or beyond. Most are from the proximal set of values.

All correlations for the benefit attitude dimensions are statistically significant for both men and women. The correlations are positive indicating attitude-intention relationships in the expected direction. They are somewhat stronger for the men than for women.

Perceived risk plays a key role in patronage intention consideration with three of the risk components being statistically significant for women and for men. As noted by the negative correlations, the higher the perceived risk, the lower is the intention to patronize a nurse-practitioner. Interestingly, the statistically significant risk components differ by sex. Social risk is a salient criteria in consideration for women while the performance risk correlation is statistically significant only for men. Financial and health risks are statistically significant correlates for both sexes.

TABLE 3  
PATRONAGE INTENTION CORRELATES

Variables	Women	Men
Symptom Sensitivity		
Psychological Readiness	.06	.16
Benefit Attitudes		
Competency	.24 <sup>a</sup>	.46 <sup>a</sup>
Interpersonal Relations	.20 <sup>b</sup>	.25 <sup>c</sup>
Treatment Performance	.35 <sup>a</sup>	.45 <sup>a</sup>
Perceived Risk		
Social	-.20 <sup>b</sup>	-.05
Financial	-.27 <sup>a</sup>	-.37 <sup>a</sup>
Health	-.17 <sup>b</sup>	-.32 <sup>a</sup>
Performance	-.11	-.32 <sup>b</sup>
Socio-demographic		
Age	-.15 <sup>c</sup>	-.02
Education	.13	.18
Children	-.01	-.12
Income	.08	.12
Health History		
Medical Scepticism	-.01	.11
Dependency in Illness	.07	.10
Medical Fatalism	.12	.24 <sup>c</sup>
Health History		
Family Doctor	-.02	.21
Recent Illness	-.03	.12
Physical Exam	-.02	.14
Recent Hospitalization	-.08	.12
Inhibiting Factors		
Physician Availability	-.10	-.22 <sup>c</sup>
Hospital Availability	-.13 <sup>a</sup>	-.20
Distance to Physician	.22 <sup>a</sup>	.15
Distance to Hospital	.07 <sup>b</sup>	.10
Insurance Plan	.19 <sup>b</sup>	-.07

- a: Statistically significant at .01  
b: Statistically significant at .05  
c: Statistically significant at .10

Among the socio-demographic variables, only age among women is statistically significant. Younger women appear more receptive to nurse-practitioners than do older women. This result is congruent with the tradi-

tional innovation literature.

Only one statistically significant relationship emerges among the health orientation concepts. This is medical fatalism. Among men, the degree of internal control correlates positively with patronage intentions.

Although none of the health history constructs are significant at the .1 level, for men the variables (especially family doctor) all point toward receptivity to this delivery system mechanism. Since men have far fewer contacts with health personnel than women (either alone or with children), the recent incident of exposure to illness and treatment may modestly increase receptivity to and trust of health care innovations.

A number of interesting relationships emerge for the inhibiting factors. For men, the availability of local health care personnel and facilities lowers the patronage intention. When contrasted to the insignificant relationships among women, this suggests a convenience orientation for men. Yet women, with perhaps an orientation toward continued medical relationships and a concern with the role of household health care coordinator, may find this factor of less importance in the consideration of alternative health care delivery methods. The physical distance to the more traditional physician seems to favor local patronage of nurse-practitioners for both sexes although it is statistically significant only for women. Finally, the availability of medical insurance relates to participation by women. Perhaps this relationship partially explains the weaker correlation for financial risk among women than among men.

The distal variables exhibit far fewer significant correlations than do the proximal variables. Since they have been frequently found to be explanatory in other health care research, it is likely that they play a moderating role in a decision process perspective taken in the current research.

## Regression Models

Separate regression models were developed for both women and men using intention as the criterion variable and the model variables as predictors. This form of analysis suffers from the assumption of direct, additive, and independent relationships among the variables involved. However, at this exploratory stage of decision model development it has major advantages. First, the cumulative influence of the predictor variables can be assessed. Second, the appearance of variables in the model that had insignificant bivariate correlations can provide clues to underlying conditional variables. Finally, should the relative influence of variables on the decision differ by sex, it will be indicated.

The regression models are given in TABLE 4. They were developed by stepwise regression methods with inclusion of variables terminated where the beta coefficients were not significant at the .1 level. The variables are illustrated by their order of entry and cumulative contribution to explained variation.

A number of observations are in order with regard to the regression models. First, both models are statistically significant ( $p < .0001$ ) and explain a moderate proportion of the variation in patronage intention. For men, the coefficient of determination,  $R^2$ , is .44, while for women the  $R^2$  is .24. It appears likely that the difference in explained variation can be attributed to the greater complexity in the women's decision, she undoubtedly evaluated the choice with a mediating force of family influences in mind.

The comparison of variables that appeared in the regression models for men and women is informative. As an

TABLE 4  
PATRONAGE INTENTION MODELS  
Stepwise Regression

Variable	Beta Coefficient	Cum. R <sup>2</sup>
<b>MEN</b>		
Intercept	-6.45	
Competency	.68 <sup>a</sup>	.21
Family Doctor	3.55 <sup>b</sup>	.25
Performance Risk	-.27 <sup>b</sup>	.30
Children	-1.31 <sup>b</sup>	.36
Education	.74 <sup>b</sup>	.40
Recent Illness	1.81 <sup>c</sup>	.44
R <sup>2</sup> = .44 F = 6.41 (p < .0001)		
<b>WOMEN</b>		
Intercept	2.61	
Treatment Performance	.29 <sup>a</sup>	.12
Distance to Physician	.19 <sup>a</sup>	.15
Social Risk	-.28 <sup>b</sup>	.18
Local Hospital	-3.26 <sup>b</sup>	.20
Local Physician	1.72 <sup>b</sup>	.22
Insurance Plan	1.11 <sup>c</sup>	.24
R <sup>2</sup> = .24 F = 7.16 (p < .0001)		

a: Statistically significant at .01  
b: Statistically significant at .05  
c: Statistically significant at .10  
n = 44 (men)  
n = 145 (women)

overview, both proximal and distal variables entered the regressions. For men, variables appeared for attitude, risk and inhibitor constructs emerged from the analysis.

In both regression models illustrated in TABLE 4, the proximal variables dominate the explanation. The salient attitude and risk constructs combine with the family doctor variable as sources of 70% of the explained variation for men. In the case of women, the salient attitude and risk constructs combine with distance to a physician to yield 75% of the explained variation.

A striking result in comparing the regressions for men versus women is that no specific construct is common to both groups. For men, performance risk is explanatory while for women the key risk variable is social risk. Likewise, with regard to attitudes, competency beliefs were dominant for men while performance beliefs appeared for women. Finally, the dominant distal variables for men are predisposition factors (children, education, illness) while inhibitors are relevant for women (distance to health care personnel, availability of local technology - facilities and insurance).

The variables that entered the regression models differed somewhat from those statistically significant through bivariate correlations. Naturally, a few of the variables with significant bivariate correlations didn't enter the equation due to moderate intercorrelations

among the predictor variables. Also, a few variables likely entered the regression as a reflection of conditional explanatory power. For women, this included the local availability of physicians and hospitals. For men, children, family doctor and education emerged in the regression models.

#### Conclusions

The major concern of the research was to investigate rural citizens' acceptance of nurse practitioners as an alternative to the present health care delivery system. A model was developed to determine the health care utilization characteristics likely to predispose one to accept nurse-practitioners as the principal primary care delivery system mechanism.

This research has shown that individuals living in rural communities are receptive to the nurse practitioner concept. The receptivity to this innovation in the health delivery system does vary within the population. The correlation analysis indicates that proximal cognitive factors (benefit attitudes and perceived risks) are quite helpful in the prediction of willingness to patronize a nurse practitioner. Thus, it appears the level of specificity of independent variables is extremely important in the prediction of patronage intent.

The multiple regression model provides an empirical assessment of the health care model with the simple framework of linear additive assumptions. The empirical analysis produces explained variances of 44% for men and 24 % for women. These are reasonably high R<sup>2</sup> values considering the exploratory nature of this research. It is interesting to note that determinant predictor variables differ by sex. For men, predictive efficiency may be greater than women ostensibly because of males' lower propensity to seek treatment of any form. Thus, men may have a simpler cognitive structure regarding health care innovations (and health care in general).

The model posited and the findings are exploratory. Yet they do merge the traditional consumer behavior theory with research in the health care field. This should serve to stimulate further consumer behavior research in the health care field and provide directions for fruitful efforts.

#### References

- L. A. Aday, The Utilization of Health Services and Indices and Correlates (Washington, D.C.: National Center for Health Services Research and Development, DHEW, Publication No. HSM 73-3003, 1972)
- L. A. Aday, Development of Indices of Access to Medical Care (Ann Arbor, MI: Health Administration Press, 1975)
- R. Andersen, A Behavioral Model of Families' Use of Health Services. Research Series, No. 25 (Chicago: Center for Health Administration Studies, University of Chicago, 1968)
- James G. Andersen and David E. Bartkus, "Choice of Medical Care: A Behavioral Model of Health and Illness Behavior," Journal of Health and Social Behavior, 14 (1973), 348-62.
- Odin W. Anderson and Paul B. Sheatsley, "Comprehensive Medical Insurance--A Study of Costs, Use and Attitudes Under Two Plans," Research Series No. 9, (Chicago: Center for Health Administration Studies, University of Chicago, 1959)
- Spyros Andreopoulos, Primary Care: Where Medicine Fails (New York: John Wiley & Sons, 1975)

- R. A. Bauer, "Consumer Behavior vs. Risk Taking," Dynamic Marketing for a Changing World(Chicago: American Marketing Association, 1960)
- Thomas W. Bice and Ken L. White, "Factors Related to the Use of Health Services: An International Comparative Study," Medical Care, 10(1969), 261-71.
- Alan Booth and Nicholas Babchuck, "Seeking Health Care from New Resources," Journal of Health and Social Behavior, 13(1972), 90-99.
- Billy Joe Franklin and Dale S. McLemore, "Factors Affecting the Choice of Medical Care among University Students" Journal of Health and Social Behavior, 2(1970), 311-319.
- Howard E. Freeman, Sol Levine, and Leo G. Reeder, Handbook of Medical Sociology, Second Ed.,(Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1972)
- R. Greertsen, "A Reexamination of Suchman's Views on Social Factors in Health Care Utilization," Journal of Health and Social Behavior, 16(1975), 226-230.
- Godfrey M. Hochbaum, Public Participation in Medical Screening Programs: A Sociopsychological Study (Washington: Public Health Service, Public Health Service Publication No. 572, United States Government Printing Office, (1958)
- John A. Howard and Jagdish N. Sheth The Theory of Buyer Behavior(New York: John Wiley & Sons, Inc.,1969)
- Stephen S. Kegeles, "Some Motives for Seeking Preventive Dental Care," Journal of the American Dental Association, 67(1963), 90-98.
- H. E. Klarman, "Effects of Prepaid Group Practice on Hospital Use," Public Health Reports, 17(1963), 955-65.
- Howard Leventhal, et al., "Epidemic Impact on the General Population in Two Cities," The Impact of Asian Influenza on Community Life: A Study in Five Cities (Washington: United States Department of HEW, Public Health Service, Publication No. 766, 1960)
- Harold S. Luft, John C. Hershey, and Joan Morrell, "Factors Affecting the Use of Physician Services in a Rural Community," American Journal of Public Health, 66(1976), 865-871.
- I. Rosenstock, "Why People Use Health Services," Milbank Memorial Fund Quarterly, 14(1966), 94-127.
- J.B. Rotter, "Generalized Expectancies for Internal Versus External Control of Reinforcement," Psychological Monographs, 80(1966)
- Edward A. Suchman, "Ethnic and Social Factors in Medical Care Orientation," Milbank Memorial Fund Quarterly, 14 (1966a), 69-78.
- \_\_\_\_\_, "Social Patterns of Illness and Medical Care," Milbank Memorial Fund Quarterly, 14(1966b), 69-78.
- Grover C. Wirick, Jr., "A Multiple Equation Model of Demand for Health Care," Health Services Research, 1 (1966), 301-46.
- Gerald Zaltman and Ilan Vertinsky, "Health Services Marketing: A Suggested Model," Journal of Marketing, 35 (1971), 19-27.
- W. G. Zikmund and Stephen J. Miller, "Rural Household Attitudes Toward Nurse Practitioners: A Factor Analytic Evaluation," Faculty Working Paper, College of Business Administration, Oklahoma State University, 1977.
- W. G. Zikmund and Jerome E. Scott, "A Multivariate Analysis of the Perceived Risk, Self-Confidence and Information Services," Advances in Consumer Research, 1(1975)



ROLE OF ATTRIBUTE GENERALITY IN COGNITIONS  
OF POLITICAL CANDIDATES

Glenn S. Omura, Michigan State University

Abstract

Following empirical developments in the study of the attribution process and voter decision making, it was proposed that image-related attributes would be more generally used than issue-related attributes to describe political candidates, but that the generality of the two types of attributes would vary by the level of office, voter party membership, and partisan or non-partisan voting behavior. The proposal was partially supported and implications to cognitive structure theory and political marketing were given.

Introduction and Background

Although political party affiliation is a major determinant of voter choice of a candidate (Campbell, 1960: 121), there is little doubt that a candidate's image and position on issues are also important voting determinants. For example, in the election of Eisenhower in 1956, and Nixon in 1972, despite the fact that over half of the voters were Democrats, both Republican presidential candidates won their elections (Agranoff, 1970: 33). The basis for the above results transcends party affiliation congruency between voters and a candidate. The purpose of this paper is to explore the theoretical roles of a candidate's image and a candidate's position on given issues in voter cognitions of the candidate.

The candidate's set of image attributes as voter determinants first attracted serious study during the 1950's, when the Eisenhower versus Stevenson campaign was researched. The positive image of Eisenhower was found to lead to greater likelihood of voting for him across party lines (Campbell *et al.*, 1954: 140). On the other hand, the Humphrey versus Nixon versus Wallace election in 1968, was found to be based on issue voting behavior among Wallace supporters, but party affiliation drew votes for Humphrey and Nixon (Converse *et al.*, 1969: 1084). Other studies show similar mixed results with regard to the relative influence of issues and images in voter choice of political candidates (Agranoff, 1970).

Although party affiliation of candidates and voters are relatively fixed, the image and issues position of candidates are not. Both the image appeal and the issues position of a candidate ideally should be congruent with those desired by a plurality of the voting public. The image of a candidate ideally should match the cognitive creation based on voter attitudes, beliefs, and values. The issue profile of a candidate ideally should match the position of the voters. Also, the image and issues strategies of a candidate must be coordinated and ought not be considered mutually independent. For example, in the Johnson versus Goldwater election, it was found that strictly on issues, Johnson was overwhelmingly favored, but his wheeler-dealer style and earlier conflict of interest charges lowered his advantage. Simultaneously, Goldwater's integrity and sincerity benefited this latter candidate, while his stand on social welfare, domestic, and foreign policy issues eroded the number of his supporters (Converse *et al.*, 1964: 331).

It is difficult to suggest more specific conclusions because past research has been relatively unsystematic. For example, the conclusions mentioned above dealt only with presidential contests. If research were to be compared across different levels of office, then stronger statements on the relative effects of image and issues might be made. A recent study by O'Keefe *et al.*, (1976) provided some data to clarify the influence of the candidate's position on the issues and the candidate's image with respect to level of office. By asking voters what was their major reason for selecting a candidate, it was found that in a presidential race, issues were mentioned more often than image characteristics. However, there was little difference between the percentage of those mentioning issues versus image reasons for the gubernatorial or senatorial races. Although situational variables may well account for the above differences, an inference may be that in higher level offices, issues are a more determinant set than image in voter decisions. Accordingly, image may be a more determinant set in lower offices. Unfortunately, the study cited above is somewhat confounded since the authors drew conclusions from the 1972 presidential race versus the 1974 gubernatorial and senatorial races.

In apparent contrast to the previous study, another recent study found that the set of image characteristics appear to be a more crucial factor than the set of issue positions in a presidential state primary election (Williams *et al.*, 1976). Williams *et al.* suggest that the reason for the greater importance of image variables is that although voters may disagree with a candidate on a few specific issues, a positive evaluation of that candidate on more general image characteristics implies a candidate's favorable performance across many situations. The apparent difference in conclusions between the O'Keefe *et al.* and Williams *et al.* studies may have been due to the fact that the former study dealt with the final election, whereas the latter study dealt with a primary election. Primaries and lower level offices may tend to create the situation which elicits voter use of general image attributes rather than specific issue attributes, possibly because of low involvement. It has been suggested by others that the degree of voter involvement varies directly with the level of political office being contested (Krugman, 1965; Yalch, 1975). Thus, the two studies may in fact have reported comparable conclusions. It was the conclusion by Williams *et al.* that image attributes may be more general than issue attributes which served as the impetus to this study.

Theory and Hypotheses

The domain of attributes relevant to the voter consists of those which the voter may ascribe or attribute to some candidate. This notion is referred to by some as differentiation. Wegner (1977) defines differentiation as "the total number of attributes given by the observer [to describe an actor]." This definition is comparable to those given by others; for example, Schroder *et al.* (1967) define differentiation as the "number of elementary dimensions in a complex cognitive structure (p. 165)." It has been noted that a characteristic of

those attributes in the differentiation domain that should be of concern in cognitive structure studies is generality. General attributes are those which are used to describe a variety of actors; where the greater the frequency of attributions, the more general the attribute. General attributes, distinct from specific attributes, have been found to be more abstract, better defined (more stable) (Wegner and Buldain, 1977), and possess more discriminating power (Wegner, 1975). Regarding the last characteristic mentioned, Schroder et al. (1967) refer to the "capacity of a dimension to distinguish stimuli" as discrimination (p. 174). They note further that one condition of an abstract attribute is the "presence of many finely separated categories on the dimension, in terms of which many aspects of stimuli can be discriminated (p. 175). To researchers of cognitive structure, they suggest a measure of discrimination developed by Scott (1962), which Wegner (1975) later used to determine that general attributes had more discriminating power.<sup>1</sup>

It is apparent from the above discussion that attribute generality, abstractness, and discrimination are positively related. This relationship provides a theoretical basis of support for Williams et al. (1976). Referring to the apparent generality of image attributes, they state that "perhaps voters approach candidate selection from an abstract perspective, making greatest use of such general attributions (p. 47)." The focus of this study was with attribute generality, relating to the suggestion by Williams et al. that perhaps image attributes are more general than issue position attributes of a political candidate. Therefore, hypotheses were developed to test whether image and issue attributes of political candidates were distinguishable on the basis of their generality. If the theoretical nature of image and issue attributes could be clarified, that is, in terms of their generality, it was felt that voter cognitions of political candidates could be better understood.

Before the hypotheses could be formally developed, however, several considerations had to be taken into account to reflect theory, past research, and the practice of political marketing. First, since the occurrence of general attributes is a function of all the attributes which may be used, both attribute characteristics of differentiation and generality had to be accounted for in the hypotheses. Second, as noted earlier, the frequency of reported usage of image versus issue attributes varied by the level of public office. Thus, level of office had to be treated as a potentially important factor. Third, as noted earlier, large numbers of voters tend to vote along party lines. Hence, party membership had to be treated as a potentially important factor, although no theoretical link between party membership and use of general attributes is immediately apparent. Also, since partisan voting is historically inconsistent, potential interacting effects between level of office and party membership had to be taken into account. Further, since partisan voting is inconsistent and there are large numbers of voters registered "independent," it was felt that understanding of voter behavior could be enhanced if voters were categorized by whether their voting behavior was partisan, non-partisan, or independent, and use of general attributes by the voter categories subsequently explored. Hence, candidates who attempt to insure partisan voting, non-partisan voting, or wish to attract independents in their campaign strategy will conceivably benefit by appropriately varying the generality of attribute descriptions. The following research hypotheses were

<sup>1</sup>Actually, Wegner used a more recent measure of discrimination developed by Scott (1969).

developed to reflect the previous discussion and the above considerations.<sup>2</sup>

- H<sub>1</sub> Frequency of usage of image attributes is greater than issue attributes at each level of office; i.e., image attributes are more general than issue attributes.
- H<sub>2</sub> Frequency of usage of image attributes is greater, relative to issue attributes, at lower level offices than at higher level offices; i.e., usage of general attributes decreases with increasing level of office relative to usage of specific attributes.<sup>3</sup>
- H<sub>3</sub> Relative frequency of usage of image/issue attributes varies with party affiliation and level of office; i.e., usage of general relative to specific attributes varies with party affiliation and level of office.
- H<sub>4</sub> Relative frequency of usage of image/issue attributes varies with patterns of voting behavior (partisan voting, crossover or non-partisan voting, and independent voting) and level of office; i.e., usage of general relative to specific attributes varies with patterns of voting and level of office.

#### Methodology

##### Operational Definitions

The approach taken to operationally define attribute generality (or specificity) was a modification of a method used in previous research (Wegner, 1977; Wegner and Buldain, 1977). The original approach was based on free response elicitation of actor attributes followed by a content analysis of the responses with attributes as the unit of analysis. In order to get a measure of a subject's use of general attributes, Wegner (1977) used a generality ratio, which was the number of attributes a subject used to describe more than one actor divided by the total number of different attributes the subject used across all given actors. Since the focus of this study was on the generality of image versus issue attributes, this concept of a generality ratio was modified in order to provide a measure of the generality of a group of attributes rather than of subjects. The operational definition used for the generality of attributes was

$$G_i = g_i/D_i$$

where:  $G_i$  = generality score for attribute category  $i$  (image or issue).

$g_i$  = number of different attributes in attribute category  $i$  used by a voter to describe all given political candidates.

$D_i$  = differentiation score or the number of different attributes in attribute category  $i$  used by a voter to describe at least one of the given candidates.

An attribute was "used by a voter to describe" a candidate if the voter could indicate a positive or negative

<sup>2</sup>The hypotheses were given alternative wordings since from a theoretical perspective, the wordings were dependent on the results of H<sub>1</sub>. Image attributes were to be termed general attributes only if H<sub>1</sub> were given empirical support.

<sup>3</sup>Note that H<sub>2</sub> is not inconsistent with H<sub>1</sub>.

assignment of that attribute. If a voter could not or did not assign an attribute to a candidate, then it was assumed that that attribute was not used to develop a cognition of the candidate.

The above operationalization provided an overall measure of the generality of image and issue attributes. Since the level of office was expected to have an effect on the use of general attributes, the above definition was modified to reflect level of office. The generality definition used for the hypotheses testing was

$$G_{ij} = g_{ij}/D_{ij}$$

where:  $G_{ij}$  = generality score for attribute category  $i$  for public office  $j$ .

$g_{ij}$  = number of different attributes in attribute category  $i$  used by a voter to describe both given political candidates contesting for office  $j$ .

$D_{ij}$  = differentiation score or the number of different attributes in attribute category  $i$  used by a voter to describe at least one of the candidates in office  $j$ .

#### Procedure and Data Collection Instrument

The data necessary to examine the hypotheses were gathered in September, 1976. Mail questionnaires were sent to a randomly selected group of registered voters in one voting district of a Boston, Massachusetts suburb. A return rate of approximately 30% yielded 335 usable questionnaires.

Voting intentions were determined by asking, "If today was election day, who would you vote for?" A number of candidates running for three offices in the selected voting district were listed. The offices were State Senator, U.S. Senator, and U.S. President. These offices were deemed to be sufficiently important to interest and involve voters, but cover a wide enough range where differences, as hypothesized between levels of office would appear.<sup>4</sup> Opposite each name, the respondent was asked to circle "YES" or "NO". No "undecided" alternative was given, but if a respondent failed to circle a "yes" or "no", it was assumed that the respondent was undecided and so classified. The names of the candidates listed that are of relevance to this study are shown in Table 1. The respondents were given no indication of the contested office or party affiliation relevant to each candidate and the candidates were presented alphabetically without regard to office sought. Voters were also asked to report their political party affiliation.

TABLE 1

#### LIST OF OFFICES AND CANDIDATES

Office	Candidate	Political Party	Incumbent or Challenger
President	Jimmy Carter	Democrat	Challenger
President	Gerald Ford	Republican	Incumbent
U.S. Senator	Edward Kennedy	Democrat	Incumbent
U.S. Senator	Michael Robertson	Republican	Challenger
State Senator	Jack Backman	Democrat	Incumbent
State Senator	Aaron Spencer	Republican	Challenger

<sup>4</sup>This assumption is later tested.

Use of an attribute by a voter to describe a candidate was determined by listing in rough alphabetical order, image and issue attributes and asking the respondent whether each specific candidate possessed the attributes. The respondent simply indicated "TRUE", "FALSE", or "DON'T KNOW". The number of "true" or "false" indications was the  $g$  value in the generality score. A "don't know" report was assumed to indicate that the respondent did not use the respective attribute in the cognitive representation of the candidate. A total of 6 image and 21 issue attributes were originally selected from the descriptions of various candidates given by the local print media and candidate promotional print brochures. In order to reduce the possibility of including trivial attributes and to balance the number of image and issue attributes in the computation of the generality scores, a decision rule was developed to select a subset of the 27 attributes. An attribute was included only if at least one-third of the voter sample responded "true" or "false" to an attribute on at least one of the two candidates in each office. The figure of one-third was arrived at by examining the raw data and determining what figure was needed to balance the number of image and issue attributes. Hence, six image and six issue attributes were finally selected and are described in Table 2.

TABLE 2

#### LIST OF IMAGE AND ISSUE ATTRIBUTES

Image Attributes	Issue Attributes
Competence	Favors pardoning draft evaders
Leadership	Favors busing
Morality	Favors handgun ban
Trustworthiness	Favors increasing dependence on foreign oil
Sincerity	Favors increased aid to elderly
Physical Appeal	Favors Equal Rights Amendment

#### Results

The first hypothesis predicted that image attributes would be found to be more general than issue attributes. In order to test this hypothesis, generality scores were computed for both categories of attributes and a series of paired sample  $t$  tests conducted. Table 3 displays the results. Regardless whether the data were collapsed

TABLE 3

#### COMPARISON OF GENERALITY SCORES

Office	Attribute Type	Mean <sup>a</sup>
Across all offices <sup>b</sup>	Image	0.110
	Issue	0.062
President <sup>b</sup>	Image	0.645
	Issue	0.548
U.S. Senator <sup>b</sup>	Image	0.313
	Issue	0.138
State Senator <sup>b</sup>	Image	0.323
	Issue	0.257

<sup>a</sup>Theoretical range is 0.0 to 1.0 where 0.0 would indicate no general attributes and 1.0 would indicate all attributes were general.

<sup>b</sup>Comparison of image and issue attribute means for the given office was significant at  $p < .01$ .

across the level of offices, or each office was examined independently of one another, the first hypothesis was supported. The generality scores for image attributes are significantly ( $p < .01$ ) larger than issue attributes.

The second hypothesis predicted that general attributes would decrease in usage with increasing levels of office relative to usage of specific attributes. This and remaining hypotheses were tested by computing a difference score from the two generality scores, that is, the dependent variable became

$$G'_j = G_{\text{image},j} - G_{\text{issue},j}$$

where:  $G'_j$  = difference score in public office  $j$ .

$G_{\text{image},j}$  = generality score for image attributes in public office  $j$ .

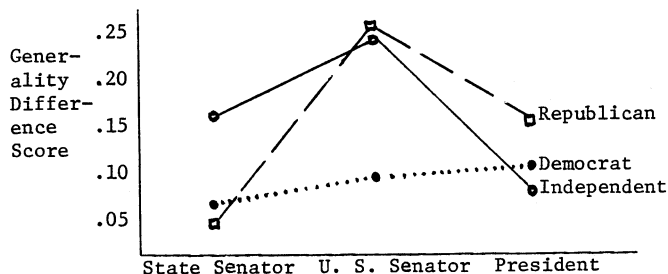
$G_{\text{issue},j}$  = generality score for issue attributes in public office  $j$ .

Since  $H_1$  was supported,  $G'_j$  was correspondingly positive. The closer the value of  $G'_j$  to zero, the closer the generality of image and issue attributes approach equality. The larger the value of  $G'_j$ , the more general were image attributes relative to issue attributes for a given level of office. This difference score reflects the hypotheses, where it is the use of the two types of attributes relative to one another in a given situation that is of concern. The statistical analysis took the form of an analysis of variance 3 x 3 within subjects repeated measures design. The two factors consisted of political party affiliation (democrat, independent, and republican) and level of office (State Senator, U.S. Senator, U.S. President). Thus,  $H_3$ , which predicted that the use of general attributes relative to specific attributes would vary by party affiliation and level of office, could be tested in the same paradigm and analysis as  $H_2$ . In fact, since it was predicted that party affiliation and level of office would have significant interaction effects, the particular ANOVA paradigm described above was required.

A significant main effect for level of office was found [ $F(2, 658) = 10.053, p < .01$ ], as well as a significant interaction effect [ $F(4, 658) = 3.159, p < .05$ ].<sup>5</sup> The main effect for party affiliation was not significant. Figure 1 describes the nature of the level of office effect and its interaction with party affiliation.

FIGURE 1

DIFFERENCES IN GENERALITY BY PARTY AFFILIATION AND LEVEL OF OFFICE



<sup>5</sup>Three subjects had to be removed from the ANOVA due to data processing limitations.

The interaction effect between level of office and party affiliation complicated conclusions about  $H_2$ . From Figure 1, it appeared that voters with a democratic affiliation were more likely to use issues as general attributes for the U.S. Senator level of office than voters with a republican or independent affiliation. Further, from Figure 1, it was evident that the graphs did not clearly suggest a decreasing function, whereas  $H_2$  predicted a monotonically decreasing function. Hence,  $H_2$  is not supported. On the other hand,  $H_3$  was partially supported since the generality scores varied with a combined (interaction) effect by party affiliation and level of office.

Since  $H_2$  was not supported, it was felt that perhaps the earlier assumption of increasing voter involvement with increasing levels of office was unwarranted. To ascertain the relationship between involvement and level of office, a simple repeated measures analysis of variance was computed for a dependent variable of "interest" and a treatment variable of levels of office.<sup>6</sup> The interest variable came from a question in the aforementioned questionnaire which asked, "How much interest do you have in the election for the following offices?" All relevant offices were listed along with a five-point response scale described by "Very Interested" to "Very Uninterested". Table 4 displays the mean interest ratings for the three offices. The analysis of variance provided some support for the earlier assumption that involvement increases with level of office [ $F(2, 334) = 121.92, p < .01$ ]. Hence, it was still unclear as to why  $H_2$  was not supported.

TABLE 4

MEAN INTEREST LEVEL BY OFFICE

Office	Mean <sup>a</sup>
President	1.916
U.S. Senator	1.961
State Senator	2.493

<sup>a</sup>Five point interest scale where Very Interested = 1 and Very Uninterested = 5.  $F(2, 334) = 121.92, p < .01$ .

The final hypothesis predicted that patterns of voting and level of office could be associated with the use of general attributes.<sup>7</sup> To test the hypothesis, a 3 x 3

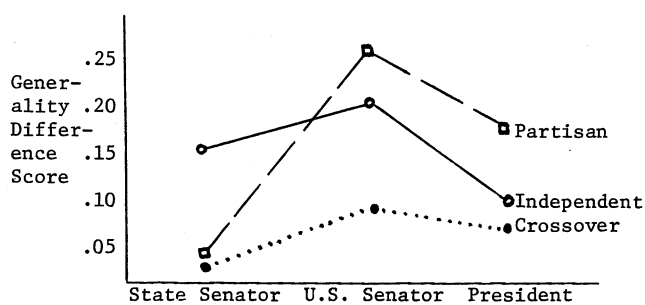
<sup>6</sup>This study is part of a larger project to study political advertising effects. The project consists of measurements over three time periods. The data used for this study came from the first questionnaire, uncontaminated by advertising treatments. However, the project allowed for several covariate measures. Hence, an "interest" variable was available for inclusion in this study.

<sup>7</sup>It is evident that the theoretical approach given in this paper implies that use of general attributes describes part of the cognitive structure of the voter. The analysis approach given for  $H_4$  implies that voting behavior affects the use of general attributes, where, in fact, no conclusion is possible on the causal direction. This analysis is therefore intended to be exploratory.

analysis of variance within subjects repeated measures design was used. The two factors consisted of patterns of voting (partisan voting, crossover or non-partisan voting, and independent voting) and the same levels of office noted earlier. Both main effects and the interaction effect were significant. The voting pattern effect was significant at  $p < .05$  [ $F(2, 329) = 4.184$ ] while both the level of office [ $F(2, 658) = 12.686$ ] and the interaction effect [ $F(4, 658) = 3.406$ ] were significant at  $p < .01$ . Figure 2 displays the results. While  $H_4$  was supported, again, because of a significant interaction effect, conclusions were not clear. The same general shape of the curves were evident when Figure 2 is compared with Figure 1. Crossover voters perhaps tended to use less generalized image attributes at lower level offices than partisan or independent voters. As was the case for  $H_3$ , the voters perhaps used more image attributes than general attributes at the U.S. Senator office level relative to the other two offices.

FIGURE 2

DIFFERENCES IN GENERALITY BY VOTING PATTERNS AND LEVEL OF OFFICE



#### Discussion

The hypotheses tested were given partial support. It was determined that image attributes were more general than issue attributes. It was found that the level of office affected use of image attributes as general attributes, but that interaction with the voter's party affiliation left conclusions somewhat ambiguous. Party affiliation *per se* did not have an effect on use of general attributes. Finally, the voting pattern of the respondents was found to vary with use of generalized attributes, but that interaction with the level of office again left conclusions unclear.

While others have suggested that general attributes are more stable and enduring, it is evident that situational considerations mediate attribute generality. In this study, for example, the level of office on which a voter must make a voting decision affected use of image versus issue attributes as general attributes. Thus, the difference in past research regarding the determinance of image attributes relative to issue attributes can be explained at least partially on the basis of the office being studied. The voting pattern results also suggest that there possibly are environmental considerations (e.g., intensity of media coverage, candidate promotion, and involvement of non-partisan advocate groups) which mediate voter generalization of attributes. Herein lies a limitation to this study (and similar survey studies): a lack of control over the immediate environment. Thus, while it may be tempting to recommend a strategy of use

or relative use of image attributes because of their generality, the results of this study could be biased by the peculiarities of the voting district used for sampling. Hence comparable research is needed in contrasting voting districts.

#### Implications

There has been much attention on the methodological aspects of cognitive and attribute modeling, e.g., ascertaining a summative or averaging cognitive procedure of processing attributes (Troutman and Shanteau, 1976); inclusion or exclusion of the evaluative aspect of the adapted Fishbein attitude model (Sheth and Talarzyk, 1972; Cohen and Athola, 1971); individual or group level of analysis (Nakanishi and Bettman, 1974); and inclusion of all or only determinant attributes (Alpert, 1972).<sup>8</sup> Selection of object attributes to label cognitive dimensions has basically evolved to the use of direct importance ratings, indirect multidimensional scaling approaches, statistical significance of belief ratings (as in regression analyses), and/or determinance (Wilkie and Pessemier, 1973; Myers and Alpert, 1968). Thus, many researchers have been concerned with whatever attributes which have been empirically proven to have the greatest weight and/or determinance without great attention to the nature of the attributes. Geistfeld et al. (1977) noted that in the past "nearly anything that a consumer perceives about a product may qualify as a product characteristic." As a more theoretically disciplined alternative, Geistfeld et al. suggested a hierarchy of product attributes with the hierarchy based on the degree of abstractness of a given product's attributes. Whether it is abstractness or the related attribute characteristic of generality focused on in this study, if theories are to be tested which involve object or person attributes, it seems theoretically prudent that researchers ought to pay more attention to the theoretical nature of attributes.

#### References

- Robert Agranoff, The Management of Election Campaigns, Boston: Holbrook Press, 1970.
- Mark I. Alpert, "Identification of Determinant Attributes: A Comparison of Methods," Journal of Marketing Research, Vol. 8, May 1972, 184-191.
- Angus Campbell, Philip E. Converse, Warren E. Miller, and Donald E. Stokes, The American Voter, New York: John Wiley, 1960.
- Angus Campbell, Gerald Gurin, and Warren E. Miller, The Voter Decides, Evanston, Ill: Row, Peterson, 1954.
- J. B. Cohen and O. T. Athola, "An Expectancy X Value Analysis of the Relationship Between Consumer Attitudes and Behavior," Proceedings, Association for Consumer Research, 1971, 344-64.
- Phillip E. Converse, Aage R. Clauson, and Warren E. Miller, "Electoral Myth and Reality: The 1964 Election," American Political Science Review, Vol. 59, 1965, 321-36.
- Phillip E. Converse, Warren E. Miller, Jerrold G. Rusk, and Arthur C. Wolfe, "Continuity and Change in American Politics: Parties and Issues in the 1968 Election," American Political Science Review, Vol. 63, 1969, 1083-1105.

<sup>8</sup>See Wilkie and Pessemier (1973) for additional issues.

Loren V. Geistfeld, George B. Sproles and Suzanne B. Badenhop, "The Concept and Measurement of a Hierarchy of Product Characteristics," William D. Perrault (ed.), Advances in Consumer Research, Vol. IV, 1977, 302-07.

H. E. Krugman, "The Impact of Television: Learning Without Involvement," Public Opinion Quarterly, 1965, Vol. 29, 349-356.

James H. Myers and Mark I. Alpert, "Determinant Buying Attitudes: Meaning and Measurement," Journal of Marketing, Vol. 32, October, 1968, 13-20.

Masao Nakanishi and James R. Bettman, "Attitude Models Revisited: An Individual Level Analysis," Journal of Consumer Research, Vol. 1, June, 1974.

Garret J. O'Keefe, Harold Mendelsohn, and Jenny Liu, "Voter Decision Making 1972 and 1974," Public Opinion Quarterly, Vol. 40, Fall, 1976, 320-30.

Harold M. Schroder, Michael J. Driver, and Siegfried Streufert, Human Information Processing, NY: Holt, Rinehart, and Winston, 1967.

W. A. Scott, "Cognitive Complexity and Cognitive Flexibility," Sociometry, Vol. 25, 1962, 405-414.

W. A. Scott, "Structure of Natural Cognitions," Journal of Personality and Social Psychology, Vol. 12, 1969, 261-272.

J. N. Sheth and W. Wayne Talarzyk, "Perceived Instrumentality and Value Importance as Determinants of Attitudes," Journal of Marketing Research, Vol. 9, 1972, 6-9.

C. Michael Troutman and James Shanteau, "Do Consumers Evaluate Products by Adding or Averaging Attribute Information?" Journal of Consumer Research, Vol. 3, September, 1976, 101-06.

Daniel M. Wegner, "The Generality and Articulation of Attributes in Person Perception," Paper presented at the Meeting of the Midwestern Psychological Association, Chicago, May 1975.

Daniel M. Wegner, "Attribute Generality: The Development and Articulation of Attributes in Person Perception," Journal of Research and Personality, Vol. 11, 1977, 329-39.

Daniel M. Wegner and Roger W. Buldain, "Attribute Generality and the Inference Process in Impression Formation," Personality and Social Psychology Bulletin, 3, 1977, 91-94.

William L. Wilkie and Edgar A. Pessemier, "Issues in Marketing's Use of Multiattribute Attitude Models," Journal of Marketing Research, Vol. 10, November, 1973, 428-41.

Daniel C. Williams, Stephen J. Weber, Gordan A. Haaland, Ronald H. Mueller, and Robert E. Craig, "Voter Decisionmaking in a Primary Election: An Evaluation of Three Models of Choice," American Journal of Political Science, 20 (February, 1976), 37-49.

Richard F. Yalch, "Attribution Theory and Voter Choice" in Schlinger, Mary Jane (ed.), Advances in Consumer Research, Vol. 2, 1975, 783-791.

## AN EXPLORATORY STUDY OF THE EVOLUTION OF THE NEGATIVE IMAGE OF PERSONAL SELLING

John J. Painter, University of Utah  
Richard J. Semelik, University of Utah

### Abstract

An exploratory study of the origins and evolution of a negative image of personal selling was undertaken among elementary, junior high and high school students using focus groups. The results suggest differences in the image of personal selling related to the age, sex, socio-economic status and past experiences of subjects.

### Introduction

A prejudice against marketing erupted several years ago, fueled by highly publicized and widely read manuscripts (Caplovitz, 1968; Packard, 1957). "Society honors those who build a better mousetrap, but suspects those who market mousetraps better." (Steiner, 1976). A more specific prejudice can be identified when personal selling as an activity and a career is considered (Crossley, S-D Surveys, 1962; Ditz, 1967; Spaulding, 1965; Smith, 1962; Staunton, 1958).

An early survey of male college students found only 1 in 17 willing to try selling (Crossley, S-D Surveys, 1962). Three-quarters of the students considered selling at best only a job, at worst a racket, while only one-quarter considered selling as a professional career. A more recent survey of high school students suggests the image of selling has not changed significantly (Sales Management, 1973).

It would appear that the image of personal selling suffers in terms of security, prestige, and financial rewards (Crossley, S-D Surveys, 1962). The image of the salesman is illustrated by a statement by Ditz (Ditz, 1967), "Being 50 years old and only a salesman reflects unfavorably both on his competence and on his status."

Personal selling, in an applied sense, has distinguished itself in our economic system. Business expenditures on personal selling are estimated at \$105 billion for 1977, and represent about three times the annual expenditures on advertising. The tremendous range of salaries in personal selling suggests there are many high level positions in the field. Many of the significantly important positions in personal selling are in industrial selling, an area to which the majority of the populace has little exposure. The marketing discipline's current view of the professional salesman is one of a "problem solver." This is a characterization that does not appear to be part of the general public's view of the profession. Consequently, the very strong possibility exists that many capable people are not considering this challenging professional opportunity since professional salespeople are suffering from an undeserved image.

### Purpose of Study

Despite the available documentation of a poor image of personal selling, there is not an identification of the origins or evolution of such an image in our society. No one would suggest the negative attitude is an inherited phenomenon, but when does it begin? When the attitude begins, its origins, dimensionality and evolution through time are useful questions to investigate. Further, as the role of women continues to change, it may

be relevant to consider a difference in attitudes toward personal selling between the female and male populations. Economic background may also contribute significantly to personal selling's image as lower economic classes may view the profession as an improvement in status (Ditz, 1967).

In order to investigate the origin, evolution and nature of the negative attitude toward personal selling, a two-stage investigation was designed. The first stage is an exploratory effort using focus group interviews to investigate attitudes among boys and girls of different ages and having different economic backgrounds. The results of this first stage are reported in this paper. The results of the current study provide useful insights regarding the influences on the attitude toward selling and its evolution as children grow up. Further, the focus group technique of the first stage provides useful insights for developing a more quantifiable investigation in the next stage of the study.

### Methodology

Focus group interviews was the method selected for this stage of the research. First, the exploratory nature of the focus group technique provided insights for further investigation of the area. Second, the desire for qualitative information, at this point, dictated an in-depth technique. Given a focus group approach to data generation, however, the generalization of results is limited (Calder, 1977).

Three criteria were considered in selecting participants for the focus groups: age, sex, and economic background. Age was considered in an attempt to gain insights regarding the point in the life cycle when a negative image of personal selling seems to emerge. Further, it provided the opportunity to identify changes in the nature of the image at different stages in the life cycle. Schools in the public school system of a large city agreed to allow their students to participate in taped discussions of career alternatives. Four age categories of students were defined by grade levels: grades 1-3, grades 4-6, grades 7-9 and grades 10-12. It was believed that students in each of these grade categories would be relatively similar in age and maturity. Participants in each focus group were drawn from a single grade category.

In order to consider a difference in attitude stemming from sex differences, half the focus group participants in each grade category were male and half were female. Each focus group, however, was restricted to strictly male or strictly female participants. The interviewer in the all male groups was male and in the all female groups the interviewer was female.

Socio-economic status was the third factor upon which the sample was partitioned. An elementary school, a junior high school and a high school in parts of the city judged to be of low socio-economic status were selected as sources for one-half of the group participants. Schools of these three types were also chosen in parts of the city judged to be high socio-economic areas and constituted the source for the other half of the group subjects. As discussed earlier, socio-economic status was

hypothesized to be a contributing or influencing factor to the image of selling. In light of this belief, the participant pool was established to isolate the variable.

Use of the three partitioning factors of age, sex, and socio-economic status resulted in sixteen different focus groups (see Figure 1). There was a total of 88 student participants in the study. No focus group consisted of less than four or more than seven subjects. Each group was purposely kept small in order to reduce the chance of individual participant intimidation due to large group size.

FIGURE 1  
Focus Group Configuration

	<u>Boys</u>		<u>Girls</u>		Total N
	Socio-economic Status		Socio-economic Status		
Grades 1-3	low (6)	high (6)	low (6)	high (4)	22
Grades 4-6	low (6)	high (6)	low (6)	high (6)	24
Grades 7-9	low (4)	high (6)	low (7)	high (5)	22
Grades 10-12	low (5)	high (6)	low (5)	high (4)	20
Total	21	24	24	19	88

Each discussion leader (interviewer) was given an identical list of questions or topics to be discussed during the taped session with each group. Discussion leaders were instructed not to vary from the prescribed order of questions or topics listed, but were permitted to probe relevant issues not included on the schedule that might arise during the session. The questions used in the focus group and their order were as follows:

- What do you want to be when you get out of school?
- What would you not want to be when you get out of school?
- What do you think marketing is?
- What do you think personal selling is?
- What kind of people do you think are in personal selling?
- How does selling compare with other jobs?
- Do you think there are different kinds of selling?
- How do you feel about the different kinds of selling?
- Why do you think you feel the way you do about selling?
- Have you ever done any selling?
- How did you feel about it?
- What does your father do for a living?
- Does your mother work? What kind of work does she do?
- What would your parents' reaction be if you became a salesperson?
- What is your overall opinion of selling and salesmen?

Each discussion leader attempted to induce participation from every subject in order to obtain opinions from each. A focus group session lasted from 45 minutes to just over one-hour.

## Results

Given the purpose of attempting to identify factors that contribute to the negative image of personal selling, results will be discussed regarding what appear to be two dominant influences: demographics and experience. Differences in the image of personal selling were related to the sex, age, and socio-economic variables used to partition the participant group. Further, the subjects' direct personal experience with selling and family experiences with salesmen also emerged as contributing factors. Finally, the type of selling task emerged as an overall consideration with regard to participant's image of the task.

### Demographics

Age. As one of the variables hypothesized to be related to the development of the image of personal selling, age exerted only a minor influence. It would appear, however, that there did exist a consistently negative progression of image as age increased. The major influence on the attitude seemed to be strongly related to personal experience. While the experience variable will be discussed later, participant responses over the four age categories suggests two influences related to age. First, it can be speculated that as the participants aged, they were more subject to direct contact with salespeople. That is, with the progression through childhood, through adolescence, to young adulthood, they became the primary target for certain personal selling efforts. When participants reached an age when they became the direct target of personal selling, most indicated they found the experience unpleasant.

A second factor related to age was participants' greater commitment to what they perceived to be the nature of the personal selling task. That is, older participants believed they had an accurate perception of what personal selling entailed and indicated that the low pay, great amount of travel, and lack of challenge in the job made it undesirable. Also related to a "knowledge" element of age, the younger participants perceive salesmen as people who "helped" other people. Older subjects did not perceive salesmen as an aid, but rather a hindrance to shopping.

Sex. The male and female participants did reveal some differences in the nature of their respective attitudes toward personal selling. Females were far more subject to parental veto of a career in personal selling. With only one exception, female participants indicated that their parents would not want them to take a job in personal selling. Conversely, only one-half of the male participants indicated that their parents would not want them to have a career in personal selling.

A second influence on the negative attitude that varied according to sex was personal experience. Female participants indicated a far greater number of positive and rewarding experiences selling than males. Citing tasks such as selling girl scout cookies, seed packets, and candy, females revealed a more pleasant reception and overall more positive experience than males. The number of such personal experiences in selling was very nearly equal between male and female subjects.

A third element that emerged related to sex was the general perception of selling. Females more often mentioned a "lack of prestige" in personal selling as a deterrent to pursuing the profession. Males, on the other hand, only rarely mentioned prestige or status and viewed poor financial rewards as the greatest drawback to personal selling. This discovery supports and lends greater clarity to the conclusions of the Crossley survey discussed earlier (Crossley, S-D Survey, 1962).

Socio-Economic Status. Partitioning the participants in



this fashion provided some interesting results. The amount of direct selling experience differed appreciably between the high and low socio-economic groups. Participants from low income areas of the city were more likely to have engaged in the childhood sorts of selling activities than their high income counterparts. Due to the negative effect of personal experience, this resulted in greater negativism among the low income participants.

Children from the low income area also perceived greater support from their parents should they choose personal selling as a career. This finding tends to support speculation that low socio-economic classes may view personal selling as an elevation in status (Ditz, 1967).

An interesting phenomenon was identified among the higher socio-economic groups relating to person-perception of those in the selling profession. Subjects frequently stated they perceived salespeople as uneducated and the type of people who could simply not find another job. Further, statements relating to dishonesty in personal selling were restricted to the higher income group participants. The precise origin of this element of negative image was hard to locate. Aside from a few experiences with high pressure door-to-door salesmen, no consistent input was identified.

#### Experiences

Direct Personal Experience. As eluded to in previous discussions, participant's direct personal experience was an important input to the image of personal selling. While the very young subjects (grades 1-3 and some subjects in grades 4-6) tended to have a reasonably positive experience selling, the effect deteriorated rapidly with age. The participants indicated several aspects of their experiences that made them unfavorably disposed toward selling. First, the exclusively door-to-door nature of their task exposed them to several unpleasant-tries: doors slammed in their faces, vicious dogs, and neighbors pretending to not be at home. Further, there were at least two elements of the reward structure that discouraged those who had tried selling. Most of their efforts were performed for very low pay. In addition to low pay (or none at all in the instance of organizational efforts), many of their selling experiences were related to "contests". That is, the person selling the most would win a skateboard or tennis racket. Since there was typically only one "winner", most of those trying personal selling found their efforts unrewarded.

There was some variability among participants' experiences with in-store sales personnel. Many of those interviewed indicated a positive and pleasant experience. Others (and the majority) related experiences whereby they felt they were ignored and/or treated rudely. The possibility exists that in-store personnel do not view young consumers as high potential prospects and therefore do not attend to them as customers as carefully as they do older shoppers.

Family Experiences. Family experiences were of two basic types: parents' or relatives' employment in personal selling and contact with door-to-door salesmen. Consistent with Ditz's findings discussed earlier (Ditz, 1967), lower income participants thought their parents or relatives who were in selling held a good job. Relative to their peers who revealed their parents were manual laborers or unemployed, personal selling (in their minds) was clearly superior. Higher income subjects (with only one exception) did not have parents employed in selling. Doctors, lawyers, and government officials dominated the employment status of the group. In light of this, the higher income subjects related more to family experiences with door-to-door salesmen.

Regardless of socio-economic status, sex, or age, parti-

cipants consistently relayed negative experiences with or attitudes toward door-to-door salesmen. The examples of a negative contact used by those interviewed could be predicted. From both personal contact and the encounters they witnessed parents having, participants indicated great distain for salesmen of vacuum cleaners, encyclopedias, magazines and other "cold call" sales approaches. Participants indicated their parents had criticism of these sales people, instructed them not to answer the door, or witnessed their parents not answering the salesman's call. Telephone sales were also cited in this category, but to a lesser degree.

#### Type of selling

Across all types of participants in the study, the image of personal selling varied consistently depending upon the type of selling discussed. In general, participants viewed real estate, automobile and in-store selling more positively. Door-to-door selling was categorically dismissed as an activity that was neither praiseworthy or feasible for consideration as a career. In conjunction, it is interesting to note that when the straightforward question "What is personal selling?" was asked, participants immediately related to the door-to-door task. This is relatively easy to explain. As stated earlier, the more sophisticated types of selling activities (industrial, wholesale, etc.) are beyond the range of experience of the general public. Since participants had little experience with two of the three types of selling viewed positively (real estate and automobile) and since the third positive type (in-store) provided them with generally negative input, the type of selling influence on image was generally negative. This, coupled with the overall negative experience with door-to-door selling, seems to provide for much of the negativism toward personal selling.

#### Summary

The variables isolated in the current study provide some meaningful inputs to the understanding of the evolution and origin of a negative image toward personal selling. In general, results of the focus group interviews suggest:

1. A progression from childhood through young adulthood provides for more negative encounters with salesmen due to lack of in-store sales assistance. Further, the older participants were subject to less rewarding personal experiences in their own selling efforts.
2. Females perceived little parental support for choosing personal selling as a career.
3. Participants from lower socio-economic backgrounds perceived more prestige in personal selling. Their counterparts from higher socio-economic backgrounds, however, downgraded the profession from a financial rewards standpoint. Further, the amount of direct experience with actually selling varied by background and contributed more negative input to the lower class' image.
4. Both family and personal experience contributed heavily to the negative image. Parents' experience with salesmen and participants' own experience was viewed negatively. One exception to the general conclusion was lower socio-economic groups' pride in a parents' or relative's employment in selling.
5. The type of selling encountered by participants contributed both negatively and positively to the image. Door-to-door selling had the greatest negative effect while real estate sales was more positively received.

The current study provided results which identify mean-

ingful elements of the negative image of personal selling. There is no attempt or especial need to try to generalize the findings of this study (Calder, 1977). Rather, the exploratory approach taken here was intended to clarify the nature of the negative image of selling and provide direction for continued research in this area.

#### References

- Bobby J. Calder, "Focus Groups and the Nature of Qualitative Marketing Research," Journal of Marketing Research, 14 (August, 1977), 353-364.
- David Caplovitz, The Poor Pay More (New York: The Free Press, 1967).
- Crossley, S-D Surveys, Inc., "Selling is a Dirty Word," Sales Management, 89 (October 5, 1962), 44-47.
- Crossley, S-D Surveys, Inc., "People Shouldn't Be Forced to Buy," Sales Management, 89 (October 19, 1962), 44-47.
- Crossley, S-D Surveys, Inc., "Salesmen are Prostitutes," Sales Management, 89 (November 2, 1962), 46-54.
- Gerhard W. Ditz, "Status Problems of the Salesman," MSU Business Topics, 2 (Winter, 1967), 68-80.
- Vance Packard, The Hidden Persuaders (New York: Pocket Books, Inc., 1957).
- Editors, "Selling? I Never Thought About It!" Sales Management, 110 (April 2, 1973), 46.
- Wendell R. Smith, "The Role of Selling in Modern Marketing," Emerging Concepts in Marketing, William S. Decker (ed.), Chicago: American Marketing Association, (Winter, 1962), 174-184.
- G.E. Spaulding, "Are We Downgrading Salesmen?" Management Review, 54 (April, 1965), 63-64.
- Donald J. Staunton, "A Survey of Public Attitudes Toward Selling as a Career," Management Review, 47 (March, 1958), 9-13.
- Robert L. Steiner, "The Prejudice Against Marketing," Journal of Marketing, 40 (July, 1976), 2-9.

AUTHOR INDEX

Olli T. Ahtola, Univ. of Florida . . . . .	253	Thomas Garman, V.P.I. . . . .	27
Gerald Albaum, Univ. of Oregon . . . . .	539	James L. Ginter, Ohio State Univ. . . . .	589
Neil Allison, Univ. of Cincinnati . . . . .	599	Marvin Goldberg, McGill Univ. . . . .	9
Mark I. Alpert, Univ. of Texas . . . . .	73, 519, 592	Linda L. Golden, Univ. of Texas . . . . .	519, 599
Pamela L. Alreck, San Diego State Univ. . . . .	139	Thomas F. Golob, General Motors Research Labs . . . . .	532
C. Dennis Anderson, Univ. of Manitoba . . . . .	488	Stephen A. Goodwin, SUNY, Buffalo . . . . .	460
David C. Arch, UCLA . . . . .	166	Marc A. Grainer, Technical Assistance Research Prog. . . . .	494
Johan Arndt, Norwegian School of Economics & Business . . . . .	128	Donald H. Granbois, Indiana Univ. . . . .	377
Stephen B. Ash, Univ. of Western Ontario . . . . .	438	Paul E. Green, Univ. of Pennsylvania . . . . .	548
Charles Atkin, Michigan State Univ. . . . .	7	Stephen A. Greyser, Harvard Univ. . . . .	501
		Kjell Grønhaug, Norwegian School of Economics & Business . . . . .	184
Richard P. Bagozzi, Univ. of California, Berkeley . . . . .	267, 295	A. William Gustafson, Texas Tech. Univ. . . . .	386
Seymour Banks, Leo Burnett Advertising . . . . .	4	Jonathan Gutman, Univ. of Southern California . . . . .	335
Sharon Banks, Univ. of Oregon . . . . .	319	Edwin C. Hackleman, Univ. of Connecticut . . . . .	90
David J. Barnaby, Univ. of Tennessee . . . . .	557	George H. Haines, Univ. of Toronto . . . . .	450
James G. Barnes, Memorial Univ., Newfoundland	550	Sherman Hanna, Kansas State Univ. . . . .	29
William O. Bearden, Univ. of South Carolina . . . . .	386	Flemming Hansen, Univ. of Copenhagen . . . . .	18
Gordon G. Bechtel, Univ. of Florida . . . . .	545	James W. Hanson, Univ. of Oregon . . . . .	622
Michael A. Belch, San Diego State Univ. . . . .	62, 139	Robert R. Harmon, Portland State Univ. . . . .	118
Russell W. Belk, Univ. of Illinois . . . . .	313	James W. Harvey, Univ. of Maryland . . . . .	209
Joseph J. Belonax, Jr., Western Michigan Univ. . . . .	232	Del I. Hawkins, Univ. of Oregon . . . . .	539
Eric N. Berkowitz, Univ. of Minnesota . . . . .	206	Roger Heeler, York Univ. . . . .	325
Roger Best, Univ. of Arizona . . . . .	539	Phillip Hendrix, Florida State Univ. . . . .	38
John F. Betak, Univ. of Texas . . . . .	519	Elizabeth C. Hirschman, New York Univ. . . . .	111, 592
James R. Bettman, UCLA . . . . .	214	Rebecca Holman, Pennsylvania State Univ. . . . .	34
Bhal J. Bhatt, SUNY, Buffalo . . . . .	460	Abraham D. Horowitz, General Motors Research Labs . . . . .	532
Stewart Bither, Pennsylvania State Univ. . . . .	353	Joel C. Huber, Duke Univ. . . . .	245, 578
Gail Bjorklund, Hood College . . . . .	400	R. Bruce Hutton, Univ. of Denver . . . . .	12
Richard Bjorklund, George Mason Univ. . . . .	400	Arun K. Jain, SUNY, Buffalo . . . . .	248
Robert C. Blattberg, Univ. of Chicago . . . . .	587	Subhash C. Jain, Univ. of Connecticut . . . . .	90
Paul N. Bloom, Univ. of Maryland . . . . .	617	Roger L. Jenkins, Univ. of Tennessee . . . . .	413
E. H. Bonfield, Temple Univ. . . . .	238	George John, Northwestern Univ. . . . .	288
Thomas V. Bonoma, Univ. of Pittsburgh . . . . .	347		
Jacques Bourgeois, Carleton Univ. . . . .	514, 550	Norman Kangun, Univ. of Houston . . . . .	184
Philip C. Burger, SUNY, Binghamton . . . . .	508	Harold H. Kassarian, UCLA . . . . .	122
Robert E. Burnkrant, Ohio State Univ. . . . .	295	K. W. Kendall, Simon Fraser Univ. . . . .	153
Alvin C. Burns, Louisiana State Univ. . . . .	372	Jerome B. Kernan, Univ. of Cincinnati . . . . .	1, 125
Robert Bushnell, Wayne State Univ. . . . .	582	Charles W. King, Purdue Univ. . . . .	20, 149
		Donald W. King, King Research . . . . .	494
Richard F. Cain, Jr., San Diego State Univ. . . . .	62	Thomas C. Kinnear, Univ. of Michigan . . . . .	38
Phillippe Cattin, Univ. of Connecticut . . . . .	284	Perri B. Koggan, Boston Univ. . . . .	66
Richard F. Chay, American Cyanamid . . . . .	259	S. Krishnan, Univ. of Pittsburgh . . . . .	445
Robert Choate, Council on Children, Media & Merchandising . . . . .	4	Zarrel V. Lambert, Auburn Univ. . . . .	611
Thomas K. Clarke, Univ. of Illinois . . . . .	313	E. Laird Landon, Jr., Univ. of Houston . . . . .	133
John D. Claxton, Univ. of British Columbia . . . . .	466	John L. Lastovicka, Temple Univ. . . . .	174
Mona Clee, Univ. of Texas, Austin . . . . .	599	Stephen A. LaTour, Univ. of Florida . . . . .	431
George Coan, Univ. of Maryland . . . . .	617	Donald R. Lehmann, Columbia Univ. . . . .	280
Joel B. Cohen, Univ. of Florida . . . . .	303	Robert P. Leone, Univ. of Texas . . . . .	245
Stanley Cohen, Advertising Age . . . . .	5	Suzanne Lindamood, Kansas State Univ. . . . .	29
Kenneth A. Coney, Arizona State Univ. . . . .	118	Jay D. Lindquist, Western Michigan Univ. . . . .	407
Fred G. Cutler, Kennedy Research . . . . .	31	William B. Locander, Univ. of Houston . . . . .	133
		Richard J. Lutz, UCLA . . . . .	329
Rachel Dardis, Univ. of Maryland . . . . .	381	Vijay Mahajan, Ohio State Univ. . . . .	248, 460
Ralph L. Day, Indiana Univ. . . . .	438	Naresh K. Malhotra, SUNY, Buffalo . . . . .	248
Margaret DeBell, Univ. of Maryland . . . . .	381	Charlene S. Martinsen, Univ. of Washington . . . . .	574
John Dimling, National Association of Broadcasters . . . . .	4	J. Barry Mason, Univ. of Alabama . . . . .	386
Ira J. Dolich, Univ. of Nebraska . . . . .	180	G. A. Mauser, Simon Fraser Univ. . . . .	161
Susan P. Douglas, Centre d'Enseignement Supérieur . . . . .	18, 364	Frederick E. May, Univ. of Missouri, St. Louis . . . . .	222
F. Robert Dwyer, Univ. of Cincinnati . . . . .	564	E. Scott Maynes, Cornell Univ. . . . .	28, 472
		Michael B. Mazis, FTC . . . . .	3
James F. Engel, Wheaton College . . . . .	169	L. G. McCabe, Consumer & Corporate Affairs, Canada . . . . .	488
John L. Evans, Consumer & Corporate Affairs, Canada . . . . .	488	James McCullough, Univ. of Arizona . . . . .	574
		Kathleen A. McEvoy, King Research . . . . .	494
Ian Fenwick, Dalhousie Univ. . . . .	153	Kenneth E. Miller, Univ. of Utah . . . . .	542
Robert Ferber, Univ. of Illinois . . . . .	146	Stephen J. Miller, Oklahoma State Univ. . . . .	628
Claes Fornell, Northwestern Univ. . . . .	105	Michael K. Mills, Univ. of Pittsburgh . . . . .	347
June Francis, York Univ. . . . .	325	Andrew A. Mitchell, Carnegie-Mellon Univ. . . . .	191
		Noreen K. Moore, Pennsylvania State Univ. . . . .	55
		Roy L. Moore, Georgia State Univ. . . . .	359

George P. Moschis, Georgia State Univ. . . . .	359, 456	Scott Ward, Harvard Univ. . . . .	7, 10, 427
Aydin Muderrisoglu, Pennsylvania State Univ. . . . .	269	Ellen Wartella, Ohio State Univ. . . . .	10
James H. Myers, Univ. of Southern California . . . . .	236, 259	Peter H. Webb, Univ. of Illinois . . . . .	85
John G. Myers, Univ. of California, Berkeley . . . . .	95	Robert A. Westbrook, Univ. of Arizona . . . . .	105
		Tracy Weston, FTC . . . . .	3
James E. Nelson, Montana State Univ. . . . .	419	Albert R. Wildt, Univ. of Florida . . . . .	569
Joseph W. Newman, Univ. of Arizona . . . . .	171	James B. Wiley, Temple Univ. . . . .	582
Larry M. Newman, Pennsylvania State Univ. . . . .	180	R. Dale Wilson, Pennsylvania State Univ. . . . .	55
		Peter L. Wright, Stanford Univ. . . . .	218
Chike Okechuku, York Univ. . . . .	325		
Richard W. Olshavsky, Indiana Univ. . . . .	392	Gerald Zaltman, Univ. of Pittsburgh . . . . .	399
Jerry C. Olson, Pennsylvania State Univ. . . . .	269	William G. Zikmund, Oklahoma State Univ. . . . .	628
Glenn S. Omura, Michigan State Univ. . . . .	635		
David J. Ortinau, Louisiana State Univ. . . . .	372		
Sadaomi Oshikawa, Univ. of Washington . . . . .	256		
John J. Painter, Univ. of Utah . . . . .	641		
Thomas L. Parkinson, Pennsylvania State Univ. . . . .	227		
Nancy C. Peat, Univ. of Florida . . . . .	431		
Steven Permut, Yale Univ. . . . .	481		
William D. Perreault, Univ. of North Carolina . . . . .	614		
J. J. Persensky, National Bureau of Standards . . . . .	15		
Edgar A. Pessemier, Purdue Univ. . . . .	606		
Ivan L. Preston, Univ. of Wisconsin . . . . .	505		
John A. Quelch, Univ. of Western Ontario . . . . .	263		
P. S. Raju, Pennsylvania State Univ. . . . .	200		
Michael L. Ray, Stanford Univ. . . . .	197		
Stanley Reid, York Univ. . . . .	325		
Michael Reilly, Pennsylvania State Univ. . . . .	227		
Peter Reingen, Univ. of South Carolina . . . . .	45		
Richard C. Reizenstein, Univ. of Tennessee . . . . .	557		
Torger Reve, Northwestern Univ. . . . .	288		
Lawrence J. Ring, Univ. of Virginia . . . . .	20		
J. R. Brent Ritchie, Univ. of Calgary . . . . .	466		
Mary Lou Roberts, Boston Univ. . . . .	66		
Thomas S. Robertson, Univ. of Pennsylvania . . . . .	10, 396		
Ivan Ross, Univ. of Minnesota . . . . .	526		
John R. Rossiter, Univ. of Pennsylvania . . . . .	424		
E. Jay Russo, Univ. of Chicago . . . . .	453		
Joel Saegert, Univ. of Texas, San Antonio . . . . .	82		
Patricia Sailor, Louisiana State Univ. . . . .	27		
Alan G. Sawyer, Ohio State Univ. . . . .	99		
Linda Scurman, Univ. of Washington . . . . .	574		
Donald Sciglimpaglia, San Diego State Univ. . . . .	62		
Carol A. Scott, UCLA . . . . .	50		
Richard J. Semenik, Univ. of Utah . . . . .	641		
Robert B. Settle, San Diego State Univ. . . . .	139		
Murphy A. Sewall, Univ. of Connecticut . . . . .	578		
Jagdish N. Sheth, Univ. of Illinois . . . . .	355		
Michael F. Smith, Indiana Univ. . . . .	392		
Theodore F. Smith, Old Dominion Univ. . . . .	392		
Rajendra K. Srivastava, Univ. of Texas . . . . .	592		
Richard Staelin, Carnegie-Mellon . . . . .	307		
Edward C. Strong, Tulane Univ. . . . .	78		
Frederick D. Sturdivant, Ohio State Univ. . . . .	529		
John L. Swasy, UCLA . . . . .	340		
James R. Taylor, Univ. of Michigan . . . . .	38		
Gerald Thain, Univ. of Wisconsin . . . . .	5		
Hans B. Thorelli, Indiana Univ. . . . .	477		
Douglas J. Tigert, Univ. of Toronto . . . . .	20, 332		
Marilyn Turner, Univ. of British Columbia . . . . .	542		
Alice Tybout, Northwestern Univ. . . . .	50		
Valerie A. Valle, Univ. of Pittsburgh . . . . .	445		
William Van Brundt, Hershey Foods Corp. . . . .	6		
Fred Van Raiij, Univ. of Tilburg . . . . .	18, 19		
M. Venkatesan, Univ. of Oregon . . . . .	34		
Alladi Venkatesh, SUNY, Binghamton . . . . .	508		
Donald E. Vinson, Univ. of Southern California . . . . .	335		
Melanie Wallendorf, Univ. of Pittsburgh . . . . .	111		
Fletcher C. Waller, General Mills, Inc. . . . .	5		
John R. Walton, Univ. of Minnesota . . . . .	206		

ORDER FORM FOR ADVANCES IN CONSUMER RESEARCH, VOLUME VI. (Ann Arbor: Association for Consumer Research, 1979).  
William L. Wilkie, Editor

International Standard Book Number: 0-915552-04-3  
International Standard Serial Number: 0098-9258

Mail Copies to (please give zip code):

Price: \$12.50 for Members of the Association for  
Consumer Research,  
\$16.00 for all others (including libraries)

Please make checks payable to the ASSOCIATION FOR  
CONSUMER RESEARCH

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
city and state: \_\_\_\_\_ zip: \_\_\_\_\_

Please send \_\_\_\_\_ copies of Advances in Consumer  
Research, Volume VI. Enclosed is \$ \_\_\_\_\_.

Please send order form to: Dr. Thomas C. Kinnear, Executive Secretary  
Graduate School of Business Administration  
University of Michigan  
Ann Arbor, Michigan 48109

---

ORDER FORM FOR ADVANCES IN CONSUMER RESEARCH, VOLUME VI. (Ann Arbor: Association for Consumer Research, 1979).  
William L. Wilkie, Editor

International Standard Book Number: 0-915552-04-3  
International Standard Serial Number: 0098-9258

Mail Copies to (please give zip code):

Price: \$12.50 for Members of the Association for  
Consumer Research,  
\$16.00 for all others (including libraries)

Please make checks payable to the ASSOCIATION FOR  
CONSUMER RESEARCH

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
city and state: \_\_\_\_\_ zip: \_\_\_\_\_

Please send \_\_\_\_\_ copies of Advances in Consumer  
Research, Volume VI. Enclosed is \$ \_\_\_\_\_.

Please send order form to: Dr. Thomas C. Kinnear, Executive Secretary  
Graduate School of Business Administration  
University of Michigan  
Ann Arbor, Michigan 48109

---

ORDER FORM FOR ADVANCES IN CONSUMER RESEARCH, VOLUME VI. (Ann Arbor: Association for Consumer Research, 1979).  
William L. Wilkie, Editor

International Standard Book Number: 0-915552-04-3  
International Standard Serial Number: 0098-9258

Mail Copies to (please give zip code):

Price: \$12.50 for Members of the Association for  
Consumer Research,  
\$16.00 for all others (including libraries)

Please make checks payable to the ASSOCIATION FOR  
CONSUMER RESEARCH

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
city and state: \_\_\_\_\_ zip: \_\_\_\_\_

Please send \_\_\_\_\_ copies of Advances in Consumer  
Research, Volume VI. Enclosed is \$ \_\_\_\_\_.

Please send order form to: Dr. Thomas C. Kinnear, Executive Secretary  
Graduate School of Business Administration  
University of Michigan  
Ann Arbor, Michigan 48109

---

IF ADDITIONAL ORDER FORMS ARE REQUIRED, PLEASE  
DUPLICATE LOCALLY

