

association for consumer research

PROCEEDINGS 3rd ANNUAL CONFERENCE 1972

M. Venkatesan, Editor



PROGRAM COMMITTEE

M. Venkatesan	The University of Iowa (Chairman)
Robert W. Pratt, Jr.	General Electric Company
Jerome B. Kernan	University of Cincinnati
Stan D. Shores	Proctor & Gamble Company
William D. Wells	University of Chicago
Gerald Zaltman	Northwestern University (Arrangements Chairman)
Fred Schlinger	Leo Burnett Company

OFFICERS OF THE ASSOCIATION (1972)

President	Joel Cohen, National Analysts Inc.
Vice President and President-Elect	Robert W. Pratt, Jr., General Electric
Treasurer	George Haines, University of Rochester
Director	Frederick Woodworth, Federal Trade Comm.
Director	William D. Wells, University of Chicago
Executive Secretary	Philip Kuehl, University of Maryland
Publications Committee	Harold H. Kassarian, UCLA, Chairman
Conference Committee (and Editor of 1972 proceedings)	M. Venkatesan, Chairman
1972 Arrangements Committee	Gerald Zaltman, Northwestern University, Chairman
Membership Committee	James Stafford, University of Houston, Chairman

PROCEEDINGS

of the

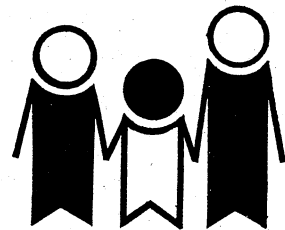
THIRD ANNUAL CONFERENCE

of the

ASSOCIATION FOR CONSUMER RESEARCH

M. Venkatesan
University of Iowa
Editor

Contributed papers presented at the
Center For Continuing Education,
University of Chicago,
Chicago, Illinois
November 3 - 5, 1972



*Copyright 1972 by the Association for Consumer Research
Library of Congress Catalog Card Number 72-97997*

Correspondence should be directed to:

Dr. Philip Kuehl, Executive Secretary
Association for Consumer Research, Inc.
College of Business and Public Administration
University of Maryland
College Park, Maryland 20742

P R E F A C E

The Third Annual Conference of the Association for Consumer Research established a milestone in the history of this association. It established a separate identity and evinced viability as an organization, even though it consists of diverse groups of people, whose focus and concern is with consumer research.

The discipline of consumer behavior is showing signs of maturity and the research efforts reflect this maturity and sophistication of methodology and analyses. Conceptual concerns are still with us and we have a long way to go in this area. All of this is reflected in the papers published in this Proceedings. Only three papers have been unacceptable for publication. Due to space limitations, neither the speeches of panel discussions are included nor the summary.

The success of this year's meeting is not only due to the Program Committee and its members and the officers of the Association, but in a large measure due to the numerous volunteers and members of the Association who actively participated in all the efforts. Particular mention must be made of those who relentlessly reviewed the research papers that were received. They are: Joel Cohen, Dave Gardner, Paul Green, George Haines, Dave Hughes, Jack Jacoby, Hal Kassarjian, Jerry Kernan, Don Lehmann, Robert Mittelstaedt, Kent Monroe, Jerry Olson, Bob Pratt, Jagdish Sheth, Stan Shores, Jim Stafford, Bill Wells, Peter Wright, and Jerry Zaltman. The untiring work of Jerry Zaltman as Arrangements Chairman, the helpful hand of Phil Kuehl, Executive Secretary, and the able assistance of John Eighmey, Graduate Assistant, all deserve special acknowledgement.

The University of Iowa
Iowa City, Iowa

M. Venkatesan
Editor

PROGRAM OF THE THIRD ANNUAL CONFERENCE

SESSIONS

Friday, November 3, 1972

11.1 PERSONALITY AND CONSUMER BEHAVIOR: The Beat Goes On . . .

Chairman: Franklin Evans, Northwestern University

Speakers: Laird Landon, University of Colorado*
A Sex-Role Explanation of Purchase Intention Differences
of Consumers Who Are High and Low in Need for
Achievement

Stewart W. Bither, Dartmouth College
Personality As A Determinant Factor in Store Choice

David W. Gardner, University of Illinois
An Exploratory Investigation of Achievement Motivation
Effects on Consumer Behavior

Thomas Kinnear, University of Western Ontario
Socioeconomic and Personality Correlates of Ecologically
Constructive Purchasing Behavior

Discussant: Masao Nakanishi, University of California at Los Angeles

11.2 CONSUMER EXPECTATIONS AND CONSUMER RESPONSE TO PRODUCTS

Chairman: James E. Stafford, University of Houston

Speakers: Rolph E. Anderson, Old Dominion University
Consumerism, Consumer Expectations, and Perceived
Product Quality

J. Taylor Sims, University of South Carolina
Consumer Response to Product-Line Extensions: A Consid-
eration for New Product Planning

Discussant: Benton Cocanougher, Southern Methodist University

*To conserve space, only the names of the first authors of joint papers
are listed.

12.1 PUBLIC POLICY REQUIREMENTS AND CONSUMER THEORIES

Chairman: Mary Gardner Jones, Federal Trade Commission

Speakers: John Howard, Columbia University
Public Policy Requirements and Consumer Theory

Robert A. Mittelstaedt, University of Missouri, Columbia
Consumer Protection and the Value of Information

Simone Clemhout, Cornell University
The Economics of "Consumerism"--Bargaining for
Resolution of Consumer Disaffection

12.2 PRICING PERCEPTIONS

Chairman: Kent B. Monroe, University of Massachusetts

Speakers: Vithala R. Rao, Cornell University
Marginal Salience of Price in Brand Evaluations

Barbara J. Deering, Purdue University
Price Intervals and Individual Price Limits as Determinants
of Product Evaluation and Selection

Jerry Olson, The Pennsylvania State University
The Relative Importance of Cues to Perceived Quality: A
Proposed Model

George J. Szybillo, Purdue University
The Relative Effects of Price, Store Image, and Composi-
tion Differences on Product Evaluation

Luncheon Speaker: Dr. George Brown, Director Bureau of the Census,
United States Department of Commerce

13.1 CONSUMER BEHAVIOR AND GOVERNMENT ACTION

Chairman: Robert Lavidge, Elrick & Lavidge, Inc., Chicago

Speakers: Ray Stokes, Consumer Research Institute Inc.
Washington, D. C.
Need to Make Greater Use of Factual Information Re-
garding Decisions Relating to Consumer Behavior

Hans B. Thorelli, Indiana University
A Concept of Consumer Policy

13.2 PERSPECTIVES ON TECHNIQUES OF ANALYSES OF CONSUMER BEHAVIOR

Chairman: Jerome B. Kernan, University of Cincinnati

Speakers: John C. Philpot, University of Tennessee
Identifying Determinants of Store Patronage Using Factor Analysis

Robert Taylor, University of Colorado
Resolving Factor Structure Distortions in Consumer Research

Flemming Hansen, T. Bak-Jensen A/S/A.I.M.,
Copenhagen
Backwards Segmentation Using Hierarchical Clustering and Q Factor Analysis

Robert Blattberg, University of Chicago
Discriminating Between Stochastic Models of Brand Choice: Minimum Chi-Square and Bayesian Methods

A. Marvin Roscoe, Jr., American Telephone & Telegraph Company
Demographic Segmentation of Long Distance Telephone Behavior

14.1 CONJOINT MEASUREMENT IN CONSUMER BEHAVIOR

Chairman: Richard Johnson, Market Facts Inc., Chicago

Speakers: John A. Fiedler, Decision Systems Group, Market Facts, Inc.
Optimizing Product Design Through Analysis of Consumer Trade Offs

Douglas Davidson, Air Canada
Forecasting Demand for a New Mode of Transportation

Paul Green, University of Pennsylvania
Consumer Menu Preference: An Application of Additive Conjoint Measurement

14.2 IN AND AROUND ATTITUDES: SOME VIEW POINTS

Chairman: S. D. Shores, Proctor & Gamble Company

Speakers: Peter L. Wright, University of Illinois
Consumer Judgment Strategies: Beyond the Compensatory Assumption

William L. Wilkie, Purdue University
Effects of the Number and Type of Attributes Included in an Attitude Model: More is Not Better

Reza Moinpour, University of Washington
An Approach to the Resolution of Multicollinearity in the Attribute Structure of Attitudes

Discussant: Thomas W. McKenna, Westat Research Inc., Rockville, Maryland

Dinner Speaker: Dr. Seymour Marshak, Marketing Research Manager, Ford Motor Company
Role of Marketing Research in Helping Management Meet the Rising Tide of Consumerism

PANEL I: On Consumer Protection Programs

Moderator: Monroe P. Friedman, Eastern Michigan University

Panelists: Helen Nelson, Director, Center for Consumer Affairs, University of Wisconsin--Extension, and President, Consumer Federation of America
A Consumer Advocate Looks at Federal Consumer Protection Programs

John Walden, Food and Drug Administration

Murray Silverman, Federal Trade Commission

Lee Richardson, Office of Consumer Affairs, Washington, D.C.

Saturday, November 4, 1972

21.1 EMPIRICAL RESEARCH AND CONSUMER PROTECTION

Chairman: Ivan Ross, University of Minnesota

Speakers: C. L. Kendall, University of North Carolina
Warranty Policies and Practices of Consumer Packaged
Goods Manufacturers

James Haefner, University of Illinois
The Perception of Deception of Advertising

Monroe Friedman, Eastern Michigan University
Consumer Response to Unit Pricing, Open Dating, and
Nutritional Labeling

Keith Hunt, The University of Iowa
Source Effects, Message Effects, and General Effects in
Counteradvertising

Discussant: Murray Silverman, Federal Trade Commission, Washing-
ton, D.C.

21.2 CONSUMER CHOICE, EXPECTATION AND RESTRICTION OF ALTERNATIVES

Chairman: Robert Witt, University of Texas at Austin

Speakers: Jacob Jacoby, Purdue University
The Components of Perceived Risk

James R. Bettman, University of California, Los Angeles
Perceived Risk: A Measurement Methodology and
Preliminary Findings

Barbara J. Deering, Purdue University
Risk Enhancement and Risk Reduction as Strategies for
Handling Perceived Risk

Michael B. Mazis, University of Florida
Consumer Reaction to Restrictions of Choice Alternatives

22.1 Ms. AND CONSUMER BEHAVIOR: SOME PERSPECTIVES

Chairperson: Marji Simon, Bozell and Jacobs, Chicago

Speakers: Michael B. Mazis, University of Florida
Attitudes Towards Women's Liberation and Perception of
Advertisements

Beverlee B. Anderson, University of Kansas
Are We Missing the Ms.? Attitudinal and Behavioral
Dimensions of Grocery Shopping

Jean Losco, University of Massachusetts
The Image of Women in Magazine Advertising

Alice Beery, Northwestern University
The Feminist Movement: Attitudes, Behavior, and Potential

Discussant: Nancy L. Thal, University of Illinois
Aleta Styres, Northwestern University

22.2 LIFE STYLE RESEARCH IN CONSUMER BEHAVIOR

Chairman: William D. Wells, University of Chicago

Speakers: Albert V. Bruno, University of Santa Clara
An Empirical Investigation of the Validity of Selected
Attitude and Activity Measures

Fred D. Reynolds, University of Georgia
An Operational Construction of Life Style

Flora L. Williams, Purdue University
Analysis of Life Styles for Study in Consumer Behavior

Discussant: Joseph Plummer, Leo Burnett Company, Chicago

Luncheon Speaker: Gerald L. Barkdoll, Asst. Commissioner, Food and
Drug Administration, Department of Health, Education
and Welfare, Washington, D.C.

23.1 CONSUMER INTENTIONS

Chairman: Robert Pratt, General Electric Company

Speakers: Donald H. Granbois, Indiana University
On the Predictive Accuracy of Subjective Purchase Probabilities

Raymond H. Suh, California State University, Long Beach
Differences Between Intenders and Non-Intenders: A Methodological Question

Discussant: C. Joseph Clawson, University of Southern California
Marketing Applications of Intentions Data

23.2 PERCEPTUAL MAPPING AND CONSUMER BEHAVIOR

Chairman: Paul Green, University of Pennsylvania

Speakers: Donald R. Lehmann, Columbia University
Buyer Behavior Models and Perceptual Mapping Models: A Synthesis

Alan B. Flaschner, University of Georgia
Comparison of Methods for Analyzing Sources of Perceived Differences in Products

Richard I. Kushner, University of New Hampshire
An Individual Differences Multidimensional Scaling Approach Towards Consumer Decision-Making

Discussant: Frank J. Carmone, Drexel University

24.1 THEORY AND THE STUDY OF CONSUMER BEHAVIOR

Chairman: Gerald Zaltman, Northwestern University

Speakers: Jagdish N. Sheth, University of Illinois
Future of Consumer Behavior Theory

David Kollat, Management Horizons, Inc.
The Current Status of Consumer Behavior Research: Developments During the 1968-1972 Period

Reinhard Angelmar, Northwestern University
Concept Validity in Consumer Behavior Theory

Discussant: Rick Pollay, Editor, Journal of Business Administration, University of British Columbia

24.2 MEDIATING FACTORS IN CONSUMER BEHAVIOR

Chairman: Charles W. King, Purdue University

Speakers: Stephen K. Keiser, University of Delaware
Social Class and Income Influences on External Search
Processes of Adolescents

Jacob Jacoby, Purdue University
Establishing the Construct Validity of Opinion Leadership

Arch G. Woodside, University of South Carolina
Dominance and Conflict in Family Purchasing Decisions

Howard Kunreuther, University of Pennsylvania
Why the Poor May Pay More for Food: Theoretical and
Empirical Evidence

James R. Bettman, University of California, Los Angeles
Social Marketing and Consumers' Preferences for Social
Consumption

PRESIDENT'S ADDRESS: Joel Cohen, National Analysts, Inc.,
Philadelphia

PANEL II: Broadening the Concept of Consumer Behavior

Moderator: Phil Kotler, Northwestern University

Panelists: Gerald Zaltman, Northwestern University

Edwin Tonneson, Bowling Green State University

Richard Coffing, University of Massachusetts

Richard Evans, University of Massachusetts

Harold Kassirjian, University of California at Los Angeles

Discussant: Dr. George Brosseau Jr., National Science Foundation

Sunday, November 5, 1972

31.1 CONSUMER SATISFACTION

Chairman: Martin Pfaff, Wayne State University

Speakers: James C. Lingo, University of Michigan
The Index of Consumer Satisfaction (ICS): Methodology

Anita B. Pfaff, Wayne State University
ICS: Results of the Pilot Test and Current Extensions

Discussants: Charles H. Handy, U.S. Department of Agriculture
Implications of the ICS for Public Policy Pertaining to
Market Performance

Robert Pratt, General Electric Company
The ICS and Corporate Marketing Policy

31.2 MEDIA AND CONSUMER BEHAVIOR

Chairman: Thomas Robertson, University of Pennsylvania

Speakers: James T. Heimback, Nationwide Research Center
The Zeigarnik Effect in Advertising

Steven W. Hollander, Standard Oil Co., Calif.
Improving Brand Recall Through "SPLIT" Commercials

George H. Haines, Jr., University of Rochester
A Study of Media Preferences and Media Consumption

John F. Willenborg, University of South Carolina
A Study of Relationship Between Social Values and
Attitudes Toward Advertising

Discussant: John G. Myers, University of California, Berkeley

32.1 NEW DIRECTIONS IN RESEARCH ON THE EXPECTANCY-VALUE ATTITUDE MODEL

Chairman: Joel Cohen, National Analysts Inc., Philadelphia

Speakers: Frederick W. Winter, University of Illinois
Mathematical Considerations in the Use of Linear Attitude
Models

(Continued on next page)

32.1 (Continued from previous page)

Alain Cousineau and Peter L. Wright, University of Illinois
Attitudes as Predictors of What?: Examining the
Dependent Variable

Richard J. Lutz, University of Illinois
Investigating the Feasibility of Personalized Rapid
Transit: An Experimental Approach

Robert E. Burnkrant, University of Illinois
Beliefs About Other Users as Predictors of Purchase
Behavior

Bobby J. Calder and Richard J. Lutz, University of Illinois
An Investigation of Some Alternatives to the Linear
Attitude Model

32.2 CROSS-CULTURAL ANALYSIS OF CONSUMER BEHAVIOR

Chairman: Flemming Hansen, T. Bak-Jensen, Copenhagen, Denmark

Speakers: Donald J. Hempel, University of Connecticut
A Cross-Cultural Analysis of Husband-Wife Roles in House
Purchase Decisions

Thomas V. Greer, University of Maryland
Cross-Cultural Considerations of Consumer Behavior: The
Case of the Consumer in the Soviet Union

Bernard Dubois, Northwestern University
A Cross-Cultural Analysis of the French Consumer

RESEARCH RESOURCES

Dr. George Brosseau Jr., Program Manager, Office of Exploratory Research
and Problems Assessment (RANN), National Science Foundation, kindly
consented to talk about the research needs and the areas of Consumer
Behavior for which NSF funding will be available. He discussed the
nature of the research program which may receive consideration for
financial support.

M. Venkatesan
Program Chairman

TABLE OF CONTENTS

PERSONALITY AND CONSUMER BEHAVIOR

A Sex-Role Explanation Of Purchase Intention Differences
Of Consumers Who Are High And Low In Need For Achievement
E. Laird Landon, Jr. 1

Personality As A Determinant Factor In Store Choice
Stewart W. Bither And Ira J. Dolich 9

An Exploratory Investigation Of Achievement Motivation Effects
On Consumer Behavior
David M. Gardner 20

Socioeconomic And Personality Characteristics As They Relate To
Ecologically-Constructive Purchasing Behavior
Thomas C. Kinnear, James R. Taylor and Sadrudin A. Ahmed . . 34

Personality And Consumer Behavior: Extensions
Masao Nakanishi. 61

CONSUMER EXPECTATIONS AND CONSUMER RESPONSE TO PRODUCTS

Consumerism, Consumer Expectations, And Perceived Product
Performance
Rolph E. Anderson and Joseph F. Hair, Jr. 67

Consumer Response To Product-Line Extensions: A Consideration
For New Product Planning
J. Taylor Sims 80

PUBLIC POLICY REQUIREMENTS AND CONSUMER THEORIES

Conceptualizing Adequacy Of Information
John A. Howard 96

Consumer Protection And The Value Of Information
Robert A. Mittelstaedt 101

The Economics Of "Consumerism": Can Collective Bargaining Work
For The Consumers?
Simone Clemhout. 107

PRICING PERCEPTIONS

Marginal Saliense Of Price In Brand Evaluations
Vithala R. Rao 125

Price Intervals And Individual Price Limits As Determinants
Of Product Evaluation and Selection
Barbara J. Deering and Jacob Jacoby 145

Cue Utilization In The Quality Perception Process
Jerry C. Olson and Jacob Jacoby 167

The Relative Effects Of Price, Store Image, And Intrinsic
Product Differences On Product Quality Evaluation
George J. Szybillo and Jacob Jacoby 180

CONSUMER BEHAVIOR AND GOVERNMENT ACTION

The Need For Problem Definition And Research Evaluation Of
Proposed Solutions For Making Public Policy Decisions Relating
To Marketing And Consumer Behavior
Raymond C. Stokes 187

A Concept Of Consumer Policy
Hans B. Thorelli 192

PERSPECTIVES ON TECHNIQUES OF ANALYSES OF CONSUMER BEHAVIOR

Identifying Determinants Of Store Patronage Using Factor
Analysis
John C. Philpot, Richard C. Reizenstein, and
Daniel J. Sweeney 201

Resolving Factor Structure Distortions In Consumer Research
Robert H. Taylor, E. Laird Landon, Jr., and
Jerome E. Scott 213

Backwards Segmentation Using Hierarchical Clustering And
Q Factor Analysis
Flemming Harsen 220

Discriminating Between Stochastic Models Of Brand Choice:
Minimum Chi-Square And Bayesian Methods
Robert Blattberg and Subrata Sen 240

Demographic Segmentation Of Long Distance Behavior: Data Analysis
And Inductive Model Building
A. Marvin Roscoe , Jr., and Jagdish N. Sheth 258

CONJOINT MEASUREMENT IN CONSUMER BEHAVIOR

Condominium Design And Pricing: A Case Study In Consumer Trade-Off Analysis
John A. Fiedler 279

Forecasting Demand For A New Mode Of Transportation
J. D. Davidson 294

Consumer Menu Preference: An Application Of Additive Conjoint Measurement
Paul E. Green, Yoram Wind, and Arun K. Jain 304

IN AND AROUND ATTITUDES: SOME VIEW POINTS

Consumer Judgment Strategies: Beyond The Compensatory Assumption
Peter L. Wright 316

Effects Of The Number And Type Of Attributes Included In An Attitude Model: More Is Not Better
William L. Wilkie and Rolf P. Weinreich 325

An Approach To The Resolution Of Multicollinearity In The Attribute Structure Of Attitudes
Reza Moinpour and James B. Wiley 341

EMPIRICAL RESEARCH AND CONSUMER PROTECTION

Warranty Policies And Practices Of Consumer Packaged Goods Manufacturers
C. L. Kendall and Frederick A. Russ 349

The Legal Versus The Behavioral Meaning Of Deception
James E. Haefner 356

Consumer Responses To Unit Pricing, Open Dating, And Nutrient Labeling
Monroe Peter Friedman 361

Source Effects, Message Effects, And General Effects In Counteradvertising
H. Keith Hunt 370

CONSUMER CHOICE, EXPECTATION AND RESTRICTION OF ALTERNATIVES

The Components Of Perceived Risk
Jacob Jacoby and Leon B. Kaplan 382

Perceived Risk: A Measurement Methodolgy And Preliminary Findings
James R. Bettman 394

Risk Enhancement And Risk Reduction As Strategies For Handling Perceived Risk
Barbara J. Deering and Jacob Jacoby 404

Consumer Reaction To Restriction Of Choice Alternatives Michael B. Mazis and Robert B. Settle	417
 MS. AND CONSUMER BEHAVIOR: SOME PERSPECTIVES	
Attitudes Toward Women's Liberation And Perception Of Advertisements Michael B. Mazis and Marilyn Beuttenmuller	428
Are We Missing The MS.? Beverlee B. Anderson	436
The Feminist Movement: Attitudes, Behavior, and Potential Alice T. Beery	446
 LIFE STYLE RESEARCH IN CONSUMER BEHAVIOR	
An Empirical Investigation Of The Validity Of Selected Attitude And Activity Measures Albert V. Bruno and Edgar A. Pessemier	456
An Operational Construction Of Life Style Fred D. Reynolds and William R. Darden	475
Analysis Of Life Styles For Study In Consumer Behavior Flora L. Williams	490
 CONSUMER INTENTIONS	
On The Predictive Accuracy Of Subjective Purchase Probabilities Donald H. Granbois and John O. Summers	502
Differences Between Intenders And Nonintenders -- A Methodological Question Raymond H. Suh	512
Marketing Applications Of Intentions Data C. Joseph Clawson	522
 PERCEPTUAL MAPPING AND CONSUMER BEHAVIOR	
Buyer Behavior Models And Attribute Models: A Synthesis Donald R. Lehmann	526
Comparison Of Methods For Analyzing Sources Of Perceived Differences In Products Alan B. Flaschner and Lyndon E. Dawson, Jr.	536
An Individual Differences Multidimensional Scaling Approach Towards Consumer Decision-Making Richard I. Kushner	546

THEORY AND THE STUDY OF CONSUMER BEHAVIOR

The Future Of Buyer Behavior Theory
Jagdish N. Sheth 562

The Current Status Of Consumer Behavior Research: Developments
During The 1968-1972 Period
David T. Kollat, Roger D. Blackwell, and James F. Engel 576

An Examination Of Concept Validity
Reinhard Angelmar, Gerald Zaltman, and
Christian Pinson 586

A Model For Research Planning In Consumer Behavior
Richard W. Follay 594

MEDIATING FACTORS IN CONSUMER BEHAVIOR

Social Class And Income Influences On External Search
Processes Of Adolescents
Stephen K. Keiser and Phillip G. Kuehl 602

Opinion Leadership And Innovativeness: Overlap And Validity
Jacob Jacoby 632

Dominance And Conflict In Family Purchasing Decisions
Arch G. Woodside 650

Why The Poor May Pay More For Food: Theoretical And Empirical
Evidence
Howard Kunreuther 660

Social Marketing And Consumers' Preferences For Social Consumption
James R. Bettman and Robert B. Andrews 679

CONSUMER SATISFACTION

The Index Of Consumer Satisfaction: Methodology
James C. Lingo and Martin Pfaff 689

An Index Of Consumer Satisfaction
Anita B. Pfaff 713

Implications Of The ICS For Public Policy Pertaining To
Market Performance
Charles H. Handy 738

The ICS And Corporate Marketing Policy
Robert W. Pratt , Jr. 742

MEDIA AND CONSUMER BEHAVIOR

- The Zeigarnik Effect In Advertising
James T. Heimbach and Jacob Jacoby 746
- A Study Of Media Preferences And Media Consumption
George H. Haines Jr., and David C. Efron 759
- A Study Of The Relationship Between Social Values And Attitudes
Toward Advertising
John F. Willenborg 783

NEW DIRECTIONS IN RESEARCH ON THE EXPECTANCY-VALUE ATTITUDE MODEL

- Mathematical Considerations In The Use Of Linear Attitude
Models
Frederick W. Winter 790
- Attitudes As Predictor Of What?: Exploring The Structure Of
Product-Related Behaviors
Alain Cousineau and Peter L. Wright 796
- Investigating The Feasibility Of Personalized Rapid Transit:
An Experimental Approach
Richard J. Lutz 800
- Beliefs About Others As Determinants Of Purchase Behavior
Robert E. Burnkrant 807
- An Investigation Of Some Alternatives To The Linear Attitude
Model
Bobby J. Calder and Richard J. Lutz 812

CROSS-CULTURAL ANALYSIS OF CONSUMER BEHAVIOR

- A Cross-Cultural Analysis Of Husband-Wife Roles In House
Purchase Decisions
Donald J. Hempel 816
- Cross-Cultural Considerations In Consumer Behavior: The Case Of
The Consumer In The Soviet Union
Thomas V. Greer 830
- A Cultural Approach To The Study Of Diffusion And Adoption Of
Innovations
Bernard Dubois 840

A SEX-ROLE EXPLANATION OF
PURCHASE INTENTION DIFFERENCES OF CONSUMERS
WHO ARE HIGH AND LOW IN NEED FOR ACHIEVEMENT¹

E. Laird Landon, Jr.²
University of Colorado

Need for achievement (n ach) as a personality variable, has received considerable attention in the psychological literature. While n ach has not been used as extensively in marketing studies, several researchers have examined its impact on various marketing phenomenon. Evans (1959), Koponen (1960), Robertson and Myers (1969), have used various measures of n ach in tandem with measures of other personality variables to explore differences between Ford and Chevrolet owners, consumer innovators and later adopters, and smokers and non-smokers. These studies have examined n ach, largely because measures have been available as part of general psychological batteries (i.e., the Edwards Personal Preference Schedule and the California Psychological Inventory). Landon (1971) used a more thorough measure of n ach to study product perception differences of high and low need for achievement consumers.

Landon's results indicate that high and low n ach consumers do differ in the way they perceive products in relation to their self- and ideal self-images. The present study presents additional findings which bear on whether or not high and low n ach consumers differ in terms of intentions to buy products and in terms of intended timing of those purchases.

Method

Of the several measures of n ach available, the Mehrabian (1968) test was administered to all subjects for the following reasons: (a) it is easy to administer and score, (b) it has a high test retest reliability, (c) it has separate scales for males and females, and (d) it appears to validly measure the need to achievement.

Each subject also rated a list of products using a five-point scale to measure purchase intentions. The first four points on the scale divided a reasonable time horizon for the purchase of the product into four divisions. For example, the four points on the scale for a carton of cigarettes were 1 week, 2 weeks, 3 weeks, and more than 3 weeks. For products purchased less often, different scales were used. The purchase of an automobile was rated on the following four points: 1 year, 2 years, 3 years, and more than 3 years. The various scales to be used with the products were each pretested to develop time horizons which were reasonable for the sample of consumers used. The fifth point on each scale (never intend to buy) was the same regardless of the product.

The products which were used for the male and female groups are presented in Tables 1 and 2, respectively. These products were chosen by five expert judges to be most likely to be related to need for achievement from a much longer list of products (Landon, 1971).

Collection and Analysis of the Data

The data used to test the hypotheses were all collected in the 1970 fall term at the University of California at Riverside. One hundred eighty-five male and 175 female freshmen and sophomores completed the questionnaire. Two forms of the questionnaire packet were used for males and two forms for females.

Within each sex, the products were presented in a randomized order and the reverse of that order.

The use of college students has made it necessary to focus the study upon consumer intentions to buy rather than more directly measuring purchase behavior. In light of the limited opportunities of college students to purchase a representative assortment of products, the researcher must either 1) limit the kinds of products to be studied, 2) measure what the subjects would like to buy, or 3) measure what they plan to buy in the future. There are always discrepancies between purchase intentions and purchase actions; however, purchase intentions may be revealing of future behavior.

Each questionnaire was placed into a high or low n ach category determined by splitting the sample at the median value of need for achievement for each sex.

Results

Ownership Aspirations

To determine whether high and low n ach consumers differ in terms of which products they plan to buy, the number of "never intend to buy" responses were tabulated and a chi-square analysis was performed. Tables 1 and 2 present the results for the male and female respondents, respectively.

In the male groups, the relationship is significant beyond the .01 level; in the female groups the relationship is not significant. While there appears to be a strong relationship between need for achievement and products in terms of the expectation of ever purchasing the product for males, the relationship is not a simple one to interpret. Of the 25 products used, 11 received more "never intend to buy" responses in the high n ach group (i.e., low n ach males will buy more of the product than will high n ach males)--while 14 of the products received more "never intend to buy" responses in the low n ach group.

Purchase Timing

Those subjects who planned, sometime in the future, to buy the products were then studied to determine whether high and low n ach consumers differ in terms of when they plan to buy the products. Tables 3 and 4 present the unequal cell size analysis of variance results. If high and low n ach consumers do indeed differ on purchase timing plans, the interaction between n ach and products should be significant. That is, on some products high n ach subjects might plan to buy sooner than low n ach subjects, but on others low n ach subjects might plan to buy sooner. The male consumers do not show a significant interaction, but the female consumers do. Table 5 shows the average time of intended purchase for each product by the high and the low n ach female respondents. Even though each product used a different time horizon, the products may be compared by interpreting low averages as indicating consumers who plan to purchase the product relatively sooner than consumers with high averages. It may be inappropriate, however, to conclude that a product with an average of 2.05 will be purchased sooner than a product with an average of 3.10.

It must be kept in mind that the results of this study indicate that need for achievement and the products measured interact in a manner which shows that purchase intentions differ. If one examines the column totals in Tables 1 and 2, it becomes apparent that with both males and females, high n ach consumers do not plan to buy more or fewer products, on the average, than do low n ach consumers.³ Comparably, with the purchase timing data, high n ach consumers do not plan to buy products on average, any sooner or later than do low n ach consumers. The nonsignificant main effect due to need for achievement in Tables 3 and 4 present this finding. However, the results of this study do

Table 1

MALE NUMBER OF "NEVER INTEND TO BUY" RESPONSES BY PRODUCT

Variable Name	Low n ach	High n ach	Total
card table and chairs	15	27	42
headache remedy	11	19	40
mouthwash	15	24	39
country club membership	53	58	111
push lawn mower	58	32	90
boating equipment	35	24	59
automatic dishwasher	13	26	39
straight razor	39	30	69
sun tan lotion	21	32	53
deodorant	6	10	16
TV dinner	18	34	52
electric toothbrush	45	59	104
coffee	33	25	58
home delivered milk	46	42	88
beer	20	17	37
imported wine	19	14	33
snow skis	42	30	72
stereo phonograph	6	2	8
art prints	24	12	36
camera	9	7	16
adult games	17	23	40
after shave lotion	19	11	30
common stock	11	20	31
dress shirt	6	4	10
travel	2	1	3
	<u>583</u>	<u>583</u>	<u>1166</u>

Chi-square = 55.6

df = 24

p < .01

Table 2

FEMALE NUMBER OF "NEVER INTEND TO BUY" RESPONSES BY PRODUCT

Variable Name	Low n ach	High n ach	Total
mouthwash	18	13	31
aluminum foil	11	4	15
headache remedy	14	10	24
automobile	2	2	4
cigarettes	77	77	154
country club membership	66	61	127
insurance	1	1	2
detergent	3	3	6
deodorant	2	1	3
color telephone	20	30	50
luggage	7	14	21
art prints	10	7	17
fabric softener	15	20	35
electric toothbrush	49	55	104
travel	0	0	0
beer	45	46	91
toothpaste	0	2	2
	<u>370</u>	<u>347</u>	<u>717</u>

Chi-square = 24.7

df = 17

p = .05

Table 3

MALE PURCHASE INTENTIONS SCORES

Analysis of Variance				
Source	SS	DF	MS	F
n ach	0,00	1	0,00	0,00
products	429,53	24	17,90	*13,50
n ach x products	29,33	24	1,22	0,92
error	4296,64	3240	1,33	

* p < .01

Table 4

FEMALE PURCHASE INTENTIONS SCORES

Analysis of Variance				
Source	SS	DF	MS	F
n ach	0,91	1	0,91	0,76
products	523,36	17	30,79	**25,64
n ach x products	34,11	17	2,01	* 1,67
error	2892,42	2409	1,20	

* p = ,05

**p < .01

Table 5
FEMALE
AVERAGE EXPECTED PURCHASE TIME

Variable Name	Low n ach	High n ach
mouthwash	2.11 70*	2.13 75
aluminum foil	3.00 77	2.96 84
headache remedy	2.51 74	2.53 78
automobile	2.97 86	3.13 86
cigarettes	2.91 11	2.09 11
country club membership	3.45 22	3.07 27
insurance	2.15 87	2.33 87
detergent	2.09 85	2.29 85
color television	3.25 81	3.38 64
deodorant	1.92 85	1.93 87
color telephone	3.40 67	3.43 58
luggage	2.82 80	2.86 74
art prints	2.03 77	1.62 81
fabric softener	3.23 73	3.24 68
electric toothbrush	3.08 39	3.03 33
travel	2.08 87	1.73 88
beer	2.88 43	3.31 42
toothpaste	2.01 87	2.02 86
	47.90 1231	47.09 1214

* cell size

suggest, that high n ach male consumers plan to buy more of some kinds of products than do low n ach male consumers. Also, high n ach female consumers do plan to buy some products sooner and some products later than do low n ach female consumers.

Discussion

There seems to be a pattern in the male data. Examining Table 1, it appears that subjects high in need for achievement tend to favor products which might be thought of as virile and masculine--such products as boating equipment, straight razor, skis, and push lawn mower. On the other hand, the male subjects scoring low in need for achievement tend to favor products which might be thought of as meticulous or fastidious--products such as automatic dishwasher, headache remedy, mouthwash, electric toothbrush, and deodorant. These differences turn out to be significant. The average probability that a high n ach subject will give a "never intend to buy" response to a "masculine" product is .32, while the probability of a low n ach male is .46 ($p < .03$). A similar analysis on the five "feminine" products reveals probabilities of .22 and .12 in the high and low n ach groups, respectively ($p < .04$).

The female purchase timing data, on the other hand, do not present any clear-cut pattern. Table 5 presents the product means for the high and low n ach females. Masculine type products such as automobiles, cigarettes, and beer are not consistently purchased sooner by either of the females. It also appears that the feminine type products will not be purchased any sooner by either high or low n ach college students.

Need for achievement is probably more closely related to the male sex role than it is to the female sex role. Men are supposed to be aggressive and competitive; but women are not. These data are consistent with this reasoning. Males who are high in need for achievement strongly identify with their sex role and project themselves via products which symbolize their sex. Females do not seem to follow this pattern. High n ach females are neither more nor less inclined toward masculine products than their low n ach counterparts. It might even be that the social stigma attached to a competitive aggressive female would force high n ach females to not express their need through purchase intentions.

If these norms toward sexual identity and need for achievement are correct, then males, in the process of internalizing norms, may come to express them through product perception and buyer behavior. The female need for achievement difference in terms of purchase timing seems not to be directly related to sex role identification.

Footnotes

1. I wish to acknowledge Harold H. Kassarian for nine-tenths of everything I know, and the UCLA campus computing network and the University of Colorado Bureau of Research for generous support.
2. The author is Assistant Professor of Marketing, School of Business, University of Colorado.
3. Interestingly enough, however, it does appear that the female subjects plan to buy more products, in total, than do the male subjects!

References

- Evans, F.B. Psychological and Objective Prediction of Brand Choice. Journal of Business, 1959, 32, 340-69.

Koponen, A. Personality Characteristics of Purchasers. Journal of Advertising Research, 1960, 1, 6-12.

Landon, E.L., Jr. Need for Achievement, Self-Concept, and Product Perception. Unpublished Phd, dissertation, University of California, Los Angeles, 1971.

Mehrabian, A. Males and Female Scales of the Tendency to Achieve. Educational and Psychological Measurement, 1968, 28, 493-502.

Robertson, T.S., & Myers, J.H. Personality Correlates of Opinion Leadership and Innovative Buyer Behavior. Journal of Marketing Research, 1969, 6, 164-168.

PERSONALITY AS A DETERMINANT FACTOR IN STORE CHOICE¹

Stewart W. Bither²
Dartmouth College
and Ira J. Dolich³

The Pennsylvania State University

In working toward a better understanding of consumer behavior, researchers have turned their attention to theories of perception, attitude formation, learning, personality and other social-psychological variables that influence the individual's decision making process. Because the monetary stakes are high in applied marketing problems, much of the research has focused on factors important in brand differentiation. Personality has not proven to be as successful a tool for predicting brand differentiation as other social psychological variables (Massy, Frank, & Lodahl, 1968). Perhaps one of the pervasive reasons for lack of success is the fact that most personality instruments used in marketing focus on interpersonal response traits. Although many brands are consumed in social situations involving more than one actor, the brand choice decision is seldom made in interpersonal social settings. Thus, for many products, individual differences of importance in interpersonal response tendencies are only indirectly related to brand choice.

There are, however, two major consumer decisions which have a direct effect on brand purchases and usually occur prior to brand choice decisions. These two decisions are the consumer's choice of product class and his choice of store. The former decision variable has received little attention in marketing although it appears, on the surface, to be an interesting and fruitful area for personality oriented research. The latter decision variable has received much attention but most of the published research focuses on demographic data and the economic analysis of store location. The present study involves an investigation of individual personality differences and their influence on store choice.

Interpersonal Aspects of Supermarket Selection

The decision to investigate the relationships between personality and store choice was predicated on the belief that store selection, although complex, involves many considerations that might be influenced by individual response characteristics of an interpersonal nature. The modern supermarket does not present the opportunity for close personal relationships between owner and patron, yet, the supermarket shopping trip is characterized by a number of interpersonal contacts. The selection and cutting of meat is often a personalized service. Check-out counters involve personal interrelationships as do check cashing situations. In addition, the occasional search for the unusual item or the item apparently out of stock often requires the help of store personnel. Finally, supermarket shopping involves maneuvering among fellow customers as well as the chance or planned meeting with friends and acquaintances. This latter characteristic has been taken into account by shopping mall planners where facilities for social activities have been incorporated in several new shopping establishments.

Despite the fact that the store choice situation is a consumer decision where present personality instruments may have predictive power, it must be acknowledged that no personality instruments have been developed specifically for the study of this, or any other consumer behavior process. In addition, previous marketing studies utilizing personality instruments have analyzed markedly different measuring instruments, consumer groups, and heterogeneous

choice situations. Thus, without a meaningful research tradition it is difficult to generalize about the potential predictability of present instruments (Kassarjian, 1971).

The personality instruments used in this study were the Edwards Personal Preference Schedule (EPPS), Jackson's Personality Research Form (PRF), and Canon's Social Adroitness Index (SAI). The EPPS was selected based on Claycamp's success with this instrument in predicting savings bank versus savings and loan patronage behavior. Using this instrument, Claycamp was able to correctly classify 72% of his subject's patronage behavior (Claycamp, 1965). Jackson's PRF and Canon's SAI were employed in this study to ascertain the applicability of newer instruments to tap individual differences relevant to marketing choice situations. These latter instruments were specifically developed for the study of nonpathological behavior and each has been validated on subjects similar to the ones used in this study.

Method

A panel of 134 housewives was established to analyze the supermarket shopping behavior of new arrivals to a community. Members of the panel were contacted prior to their arrival to the new community. Cooperation in pre-arrival identification of prospective newcomers was obtained from industry, the Chamber of Commerce, real estate agencies and colleges located in the area. Pre-arrival shopping factors were measured through a mail questionnaire and post-arrival data was obtained by personal interview three months after arrival. During the post-arrival interview, subjects completed the Edwards EPPS, the Jackson PRF, the Canon SAI, and a comprehensive questionnaire about their supermarket shopping since their arrival.

Panelists were classified into three groups according to their usage of a most preferred supermarket (Farley, 1968). Those women who made 60% or more of their shopping trips to their most preferred supermarket were classified as regular shoppers. Irregular shoppers were defined as those women who made less than 40% of their shopping trips to their most preferred supermarket. The remainder of the panel were designated as intermediate shoppers. Since the major thrust of the investigation was to study regular versus irregular shoppers, the intermediate shoppers were set aside. The entire battery of personality traits and a standard collection of socioeconomic variables were then utilized as independent variables in building discriminant functions.

A stepwise discriminant analysis was utilized to establish the functions (Dixon, 1971). The stepwise procedures were used because of the exploratory nature of this study. Although the EPPS had been used in previous research (cf. Claycamp, 1965; Evans, 1959; Keuhn, 1963), insufficient evidence existed to identify relevant traits for supermarket choice behavior on an a priori basis. Further, the Canon SAI had not been previously used in marketing research and the Jackson PRF had not been used in studies of this kind.

The Jackson and Canon instruments were separated from the Edwards Schedule to illustrate the discriminating ability of the two sets of traits individually as well as in combination. This was a procedure utilized by Claycamp to indicate relative importance of classes of variables. Separation of trait instruments in this study is not intended to indicate importance, but to show the loss of discriminating power of individual instruments relative to combined trait usage.

Limitations

The results of this research project are subject to the following limitations:

1. The community in which the study was conducted is dominated by a large university. Thus, the panel of newcomers differed from average United States demographic characteristics on level of education.
2. The panel differed from established adult norms on several traits of the personality instruments used.
3. The study included store choice behavior only during the first three months of a newcomers arrival.
4. The community studied is a medium sized community with approximately 12 full line supermarkets.

All of the above factors limit the extent to which the results of this study may be generalized.

Results

There are numerous studies which have reported improved explanation of regression models and increased classification in discriminant models when socioeconomic variables were "tossed into the pot." Demographic variables were also included in the present analysis. However, the stepwise procedures did not provide socioeconomic variable usage meeting several important criteria: (1) parsimonious use of variables, (2) useful increases in proper classifications, and (3) significant differences in discriminant mean values. Thus, the discriminant functions were not improved by adding socioeconomic variables. It should be noted, however, that the newcomer panel did not have substantial variation on the demographic variables of income, education, and age.

The Multiple Discriminant Analysis of Regular and Irregular Shoppers

The 41 regular and 39 irregular shoppers were randomly divided into split half samples. Subsample 1 utilized 21 regular and 20 irregular subjects and subsample 2 contained 20 regular and 19 irregular shoppers. Each discriminant function was derived from subsample 1 and tested in subsample 2 (cf. Wilson, Mathews & Sweeney, 1971; Cooley & Lohnes, 1962; Overall & Klett, 1971; Frank, Massy & Morrison, 1965; Morrison, 1969). A total of five discriminant models were derived and tested. Model 1 consists of the best possible model using only the EPPS traits to predict store patronage. The second model uses the combined traits of the Jackson PRF and Canon SAI with store patronage as the dependent variable. The third model uses all three personality instruments as independent variables with patronage behavior as the dependent variable. Model 4 indicates the discriminating ability of the EPPS variables, and Model 5 indicates the discriminating ability of the PRF and SAI variables used in Model 3. Table 1 shows the traits included in each of the 5 models as well as the canonical coefficients and F tests for each of the traits. Table 2 shows the classification matrices for each of the five models for the subsample from which the model was derived and for the subsample on which the model was tested. Table 3 is a summary of the percentage of properly

classified cases by each of the 5 models and Table 4 shows the approximate F tests for each of the 5 models.

Model 1. Model 1, using the EPPS, predicts 85 percent correct classification (Tables 2 & 3), using the traits of achievement, deference, autonomy, succorance, dominance, endurance, and aggression. However, validation with subsample 2 provides only 54 percent correct assignments. This occurs even with the strong significance test results ($P[F] = < .01$) of actual differences existing between the two discriminant means (Table 4). It is interesting to note that Claycamp found (among 6 discriminating traits) the three traits of achievement, deference, and autonomy in his stepwise discriminant models of savings depositors (Claycamp, 1965). And, Evans (Evans, 1959) used achievement, deference, autonomy, dominance, aggression, and six other traits in his often quoted study of Ford and Chevrolet owners.

Model 2. The best possible classification model utilizing the PRF and SAI is shown as Model 2 (with 6 traits). This model suffered the same fate as Model 1 in terms of precipitous drops in classification ability for the test group. Tables 1 through 4 provide most of the detail necessary for further evaluation.

Model 3. Model 3, with 5 traits, is the most parsimonious configuration obtained for the entire set of traits. No improvement in the 78 percent correct classification is obtained until the addition of a 15th trait. Testing the model against subsample 2 provides 70 percent correct classification, the best obtained for any model. The F values of the variables are all in reasonable ranges (Table 1) as well as approximate F test for Wilks Lambda (Table 4). A summary table of mean values is also provided as Table 5. Figure 1 provides plots of subsample 1 and Figure 2 provides plots for subsample 2 based on Model 3 coefficients.

The combined instrument model utilizes 2 EPPS (dominance and intraception), 2 PRF (social approval and order), and the SAI scale. It is interesting to note that Evans and Claycamp (Claycamp, 1965; Evans, 1959) both used intraception and dominance but the latter eliminated intraception in his stepwise procedures. It would appear that the EPPS Dominance trait has relevance in marketing across several behavioral measures. Claycamp's best results were with 6 personality variables alone (autonomy, achievement, recession, nurturance, deference, and heterosexuality) where 79 percent were properly classified. However, he did not have a test of the model to determine the magnitude of the upward bias. It should be noted that Model 1, a model with 7 EPPS traits and significant F at .01, rather than Claycamp's .05, can drop over 30 percent in correct assignment on testing with a split-half sample.

Models 4 and 5. Models 4 and 5 are provided to show the relative capabilities and characteristics of the individual sets of personality traits. Although neither model performs as well as Model 3, they do not fair badly nor indicate the presence of the large upward biases found with the more numerous trait functions of Models 1 and 2.

Table 1
Subsample 1 Discriminant Functions^a

Model 1:		Canonical	F	d.f.	P(F)
<u>EPPS Traits</u>		<u>Coefficients</u>			
1	Achievement	-0.066	0.786	1, 33	-
2	Deference	-0.028	10.091		<.01
5	Autonomy	-0.479	18.842		<.01
8	Succorance	0.157	3.083		.10
9	Dominance	-0.065	1.116		.30
13	Endurance	-0.117	1.571		.25
15	Aggression	0.057	0.924		-
Model 2:					
<u>PRF/SAI</u>					
2	Social Approval	-0.251	3.539	1, 34	.10
4	Dominance	-0.037	0.139		-
6	Order	-0.138	2.196		.20
7	Affiliation	-0.112	0.940		-
8	Achievement	-0.048	0.322		-
1	Social Adroitness	-0.148	2.915		.10
Model 3:					
<u>EPPS/PRF/SAI</u>					
7	Intracception	-0.198	3.030	1, 35	.10
9	Dominance	0.152	6.523		.05
2	Social Approval	0.219	6.560		.05
6	Order	0.153	4.796		.05
1	Social Adroitness	0.099	2.536		.15
Model 4:					
<u>EPPS</u>					
7	Intracception	-0.156	.615	1, 38	-
9	Dominance	0.182	3.196		.10
Model 5:					
<u>PRF/SAI</u>					
2	Social Approval	.200	3.131	1, 37	.10
6	Order	.114	1.551		.25
1	Social Adroitness	.138	2.579		.15

^aBMD07M

Table 2

Classification Matrices of 5 Discriminant Models From
Subsample 1 and the Test Against Subsample 2

	Discriminant Model				Test Group			
	Actual	Assigned		Tot	Actual	Assigned		Tot
1. EPPS 7 Traits	I	18	2	20	I	10	9	19
	R	<u>4</u>	<u>17</u>	<u>21</u>	R	<u>9</u>	<u>11</u>	<u>20</u>
		22	19	41		19	20	39
2. PRF/SAI 6 Traits	I	12	8	20	I	13	6	19
	R	<u>5</u>	<u>16</u>	<u>21</u>	R	<u>15</u>	<u>4</u>	<u>20</u>
		17	24	41		28	10	39
3. EPPS/PRF/SAI 5 Traits	I	16	4	20	I	15	4	19
	R	<u>5</u>	<u>16</u>	<u>21</u>	R	<u>8</u>	<u>12</u>	<u>20</u>
		21	20	41		23	16	39
4. EPPS 2 Traits	I	11	9	20	I	12	7	19
	R	<u>7</u>	<u>14</u>	<u>21</u>	R	<u>8</u>	<u>12</u>	<u>20</u>
		18	23	41		20	19	39
5. PRF/SAI 3 Traits	I	12	8	20	I	13	6	19
	R	<u>7</u>	<u>14</u>	<u>21</u>	R	<u>8</u>	<u>12</u>	<u>20</u>
		19	22	41		21	18	39

Table 3

Summary of Proper Classification for 5 Discriminant Models
from Subsample 1 Tested with Subsample 2

	Subsample	
	1	2
1. EPPS (7)	85%	54%
2. PRF/SAI (6)	72%	44%
3. EPPS/PRF/SAI (5)	78%	70%
4. EPPS (2)	61%	62%
5. PRF/SAI (3)	64%	64%

Table 4
Approximate F Statistics for 5 Discriminant Models^a

1		d.f.	F	P[F]
1.	EPPS 7 Traits	7, 33	3.90	.01
2.	PRF/SAI 6 Traits	6, 34	1.13	.30
3.	EPPS/PRF/SAI 5 Traits	5, 35	2.89	.05
4.	EPPS 2 Traits	2, 38	1.64	.25
5.	PRF/SAI 3 Traits	3, 37	2.03	.25

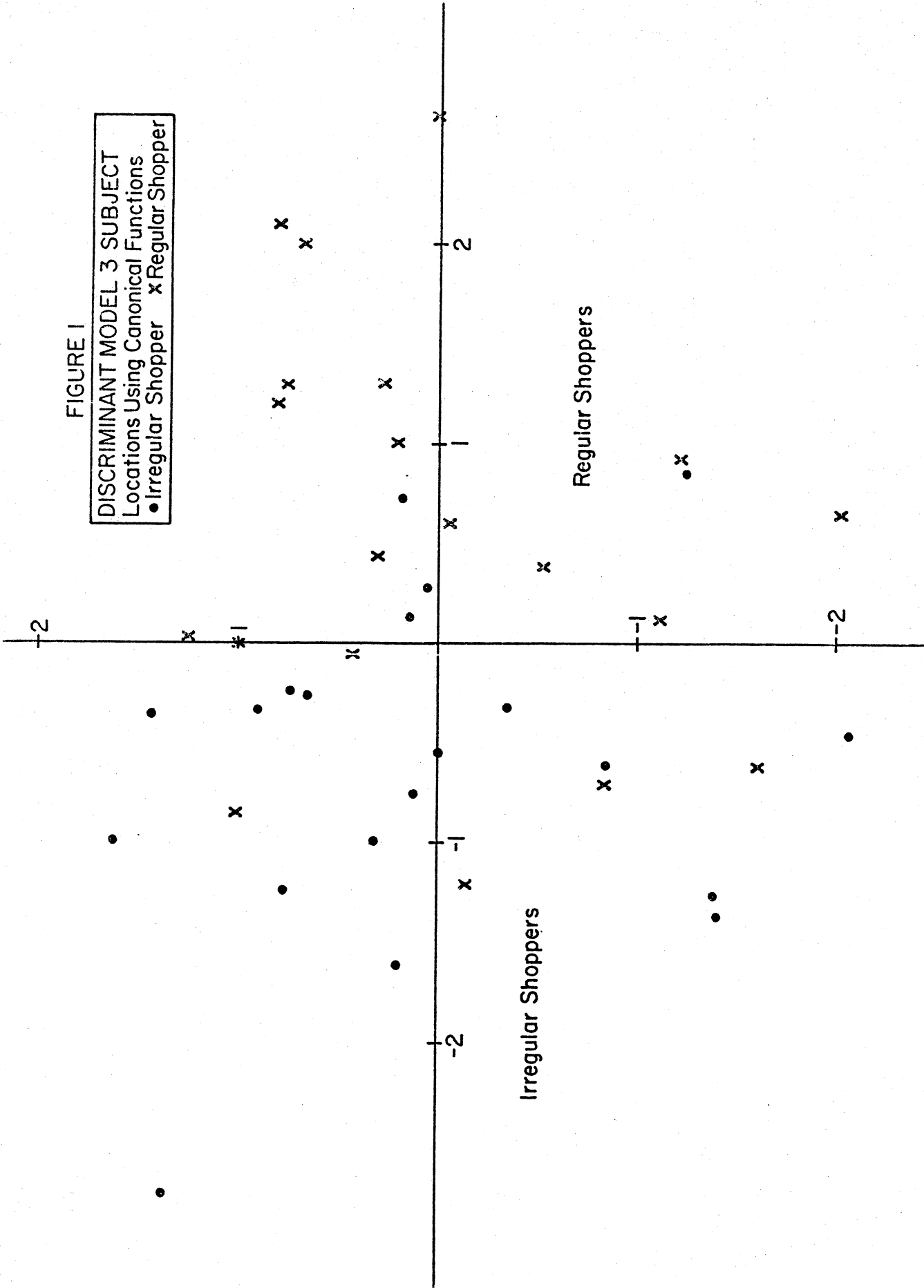
^aBMD07M

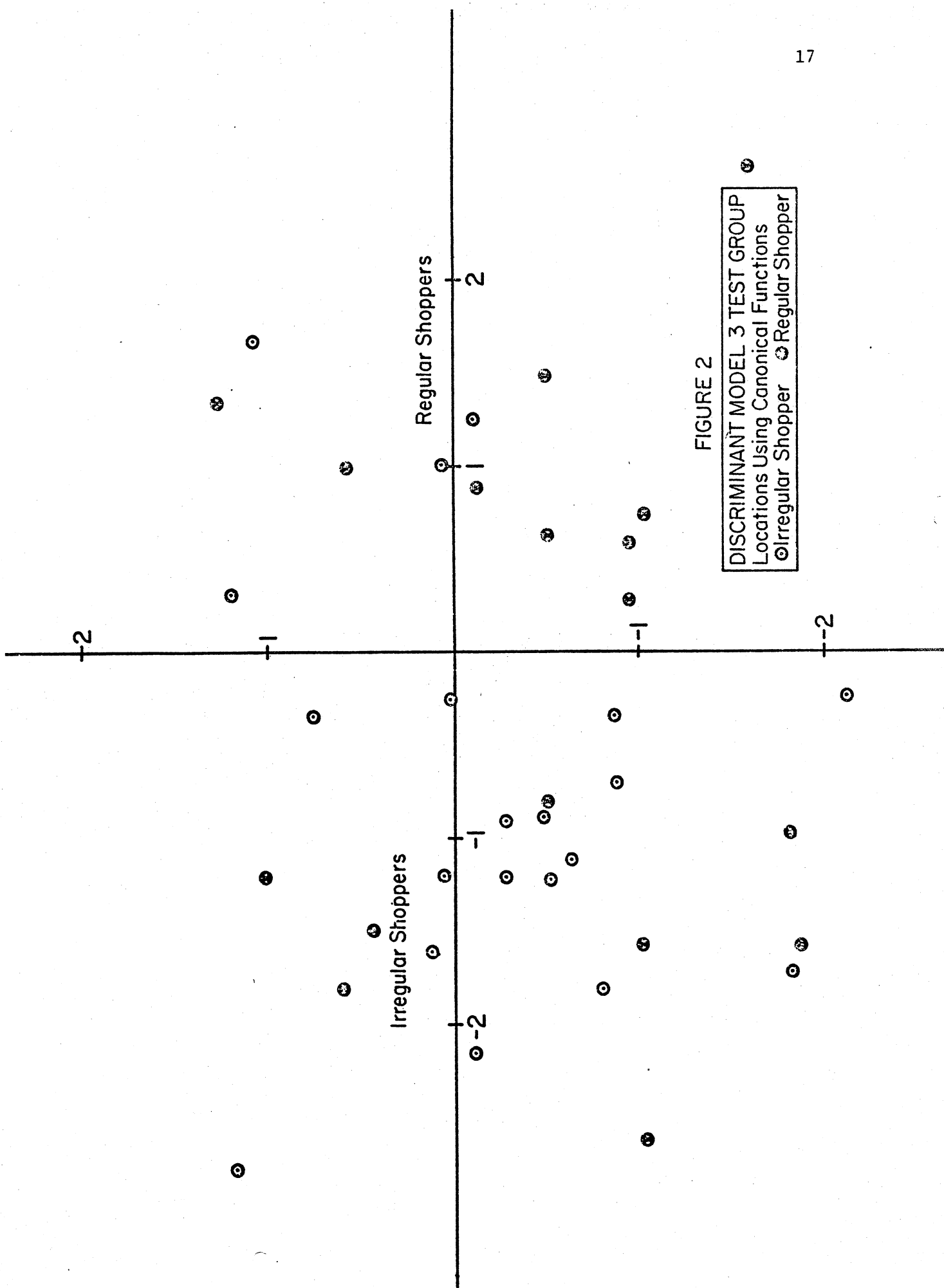
Table 5

Summary Table of Model 3 Trait Means

	Subsample 1 Means		Subsample 2 Means		Total Sample Means
	Regular	Irregular	Regular	Irregular	
EPPS					
Intraception	15.10	15.35	13.80	15.74	14.99
Dominance	23.52	-20.60	22.30	21.58	22.02
PRF					
Social Approval	13.66	11.80	12.15	11.42	12.29
Order	20.76	20.05	18.35	19.05	19.58
SAI					
Social Adroit- ness	12.14	9.70	11.90	10.16	11.00

FIGURE I
DISCRIMINANT MODEL 3 SUBJECT
Locations Using Canonical Functions
● Irregular Shopper x Regular Shopper





DISCRIMINANT MODEL 3 TEST GROUP
Locations Using Canonical Functions
○ Irregular Shopper ● Regular Shopper

FIGURE 2

Summary

This study involved an investigation of the relationship between individual personality differences and store patronage behavior. A rationale for investigating the role of personality in this area of consumer decision making was developed based on the interpersonal aspects of the consumer shopping trip. A panel of 134 newcomers to a community participated in the study. Regular and irregular usage of the consumers most preferred store was used as the dependent measure of patronage. Independent variables were the individual trait scales of 3 personality instruments, the Edwards Personal Profile Schedule, Jackson's Personality Research Form, and Canon's Social Adroitness Index. Discriminant models were developed using these instruments singly and in combination. Each model was derived and tested on separate halves of the panel data.

The most parsimonious and predictive of the models developed used a combination of traits from the three personality instruments. Seventy percent of the subjects were properly classified in the test of this model. Although use of the Edwards instrument in consumer studies has been reported, two of the personality instruments have seen little use in marketing research (the Canon instrument has not been previously used). Some similarity between the present study and earlier studies in the prediction ability of individual traits on this instrument was observed.

Footnotes

1. This research was supported by a National Science Foundation institutional grant to the two authors by The Pennsylvania State University.
2. Stewart W. Bither is Visiting Associate Professor at the Amos Tuck School of Business Administration, Dartmouth College and Associate Professor of Marketing at The Pennsylvania State University.
3. Ira J. Dolich is Associate Professor of Marketing at The Pennsylvania State University.

References

- Claycamp, H. J. Characteristics of owners of thrift deposits in commercial banks and savings and loan associations. Journal of Marketing Research, 1965, 2, 163-170.
- Cooley, W. W. & Lohnes, P. R. Multivariate Procedures in the Behavioral Sciences, New York: John Wiley, 1962.
- Dixon, W. J. (Ed.), BMD: Biomedical Computer Programs, Berkeley: University of California Press, 1971.
- Evans, F. B. Psychological and objective factors in the prediction of brand choice. Journal of Business, 1959, 32, 340-369.
- Farley, J. U. Dimensions of supermarket choice patterns. Journal of Marketing Research, 1968, 5, 206-208.

- Frank, R. E., Massy, W. F., & Morrison, D. G. Bias in multiple discriminant analysis. Journal of Marketing Research, 1965, 2, 250-258.
- Kassarjian, H. H. Personality and consumer behavior: a review. Journal of Marketing Research, 1971, 8, 408-418.
- Keuhn, A. A. Demonstration of a relationship between psychological factors and brand choice. Journal of Business, 1963, 36, 237-241.
- Massy, W. F., Frank, R. E., & Lodahl, T. Purchasing behavior and personal attributes, Philadelphia: University of Pennsylvania Press, 1968 and Journal of Advertising Research, 1969, 9.
- Morrison, D. G. On the interpretation of discriminant analysis. Journal of Marketing Research, 1969, 6, 156-163.
- Overall, J. E. & Klett, C. J. Applied multivariate analysis. New York: McGraw-Hill Book Company, 1971.
- Wilson, D. T., Mathews, H. L. & Sweeney, T. W., Industrial buyer segmentation: a psychographic approach. Presented at The American Marketing Association's Fall Conference in Minneapolis, Minnesota, August 30 - September 1, 1971.

AN EXPLORATORY INVESTIGATION OF
ACHIEVEMENT MOTIVATION EFFECTS ON
CONSUMER BEHAVIOR

David M. Gardner
University of Illinois at
Urbana-Champaign
and Federal Trade Commission
Office of Policy Planning and Evaluation

Not too many years ago, "motivation research" was a widely used term and technique to explore certain aspects of consumer behavior. The term and technique are no longer in vogue because many researchers feel that the global claims of motivation research glossed over effects of many explanatory variables. The term "motivation" is now referred to by most researchers as an important construct in explaining consumer behavior, but only one of several, i.e., attitudes, learning. But while many writers recognize that theories of motivation can be used in the explanation of consumer behavior, the literature is deficient in relating motivation (especially learned motivation) variables to consumer behavior.

The purpose of the study reported in this paper is to explore the relationship of achievement motivation to the assortment of goods and services held and desired by consumers. Underlying this study is the basic question: Does need for achievement explain differences in purchasing behavior; and if so, does it deserve further research?

The study of motivation is concerned with the needs, wants and fears of man, and how he seeks to attain or avoid these things. Motivation is the study of activation and direction of behavior. The reason that there has been little actual research or application of motivation to consumer behavior stems in part from two problems. Motivation researchers are themselves divided on some basic issues. These differences are apt to be amplified when adapted by another discipline. Secondly, learning theories and motivational theories are highly intertwined with each other, which makes it difficult for the outsider to use and understand them.

The study of motivation is complicated by the existence of two major theories. The first is the well-known "Drive X Habit" theory, so often associated with Hull (Atkinson, 1964). In this conception of motivation, drive (a non-specific energizer) is said to combine with learned habits to activate behavior when associated with a relevant stimulus situation. Drive is the product of some need, want, or fear, which results in the individual being energized to satisfy that need, want, or fear. Drive combines with habit strength, which is a product of the frequency of prior need reducing states. Individuals develop habits which become associated with stimulus situations. Drive is said to provide the activation in a stimulus situation, habit the direction of behavior.

We can say then, that when faced with a stimulus situation, a person will respond in a given manner, depending upon the habit associated with similar prior circumstances and the amount of drive present in the system.

Therefore, a person presented with a well-prepared meal may not eat the meal, if his drive level is low, as a result of having eaten a large meal. On the other hand, if a person is almost completely satiated with foods, the habit strength of some foods may be so strong so as to result in eating of those foods, even though drive is quite low.

A major deficiency in the "Drive X Habit" theory is the lack of concern with the consequence of an act. The individual's motivation is not tempered or heightened by any type of expectations. The second major theory, "Expectancy X Value," avoids this problem. This theory, no matter who the advocate, suggests the "strength of the tendency to act in a certain way depends upon the strength of expectancy that the act will be followed by a given consequence (or goal), and the value of that consequence (or goal) to the individual (Atkinson, 1964).

A specific formulation of the "Expectancy X Value" theory is that of achievement motivation. This formulation suggests that: "strength of the tendency to act in a certain way to attain a particular goal in a particular situation is influenced by a relatively non-specific variable called motive (M), which is tentatively assumed to be a relatively stable characteristic of a person carried about from one life situation to another, and two relatively specific influences which refer to the particular act in question and which are defined by cues in the immediate environment: the strength of expectancy that the act will be followed by a particular consequence (E) and the incentive value of that particular consequence (I) (Atkinson, 1964).

More specifically, this theory suggests that people develop at an early age (before the age of ten) a tendency to achieve. People with high levels of this tendency (High n Achievement) have been shown to be concerned about success, tend to engage in achievement-related activities, and to prefer tasks of intermediate difficulty (Atkinson, 1964). People with low levels of this tendency (Low n Achievement) have been shown to be concerned about failure, tend to avoid achievement-related tasks, and prefer tasks which are too easy or too difficult in relation to the person's abilities. Atkinson (Atkinson, 1964) has postulated that this tendency to achieve (M_s) combines multiplicatively with situation influences to produce a tendency to approach success. For Atkinson, the situational factors are: the strength of expectancy or probability of success (P_s) and the incentive value of success of a particular activity (I_s). The tendency to approach success (T_s) or the strength of the motive to achieve is represented by Atkinson as: $T_s = M_s \times P_s \times I_s$. "The first variable, M_s is a relatively general and stable characteristic of the person which is present in any behavior situation. But, the values of the other two variables, P_s and I_s , depend upon the individual's past experience in specific situations that are similar to the one he now confronts. These variables change as the individual moves about from one life's situation to another and so are treated as characteristics of particular situations or particular tasks" (Atkinson, 1964).

Related Research

Since McClelland advanced the general theory of achievement motivation and its measurement (McClelland, Atkinson, Clark, & Lowell, 1953), a number of studies have been conducted to identify the relationship

between levels of n-achievement and other psychological and performance variables. While over one-hundred articles appear in the literature, several seem especially relevant as background for this study.

Atkinson suggests (Atkinson, 1957) that people with a high level of need of achievement desire to be in situations of moderate risk as opposed to situations of either low or high risk preferred by people with low levels of need for achievement. Littig (Littig, 1959) and Litwin (Litwin, 1958) have demonstrated this preference, using game situations. McClelland and Liberman found that high need achievers recognized positive achievement words more quickly than middle and low need achievement groups (McClelland & Liberman, 1949).

Knapp has shown that people with high levels of need for achievement prefer subdued bluish tartans, while a preference for red and bright tartan plaids is exhibited by people with low levels of need for achievement (Knapp, 1958). Using the A0 scale of the California Personality Inventory, Carney has shown that achievement motivation and smoking are positively correlated (Carney, 1967).

In addition to these studies, there are consistent findings that males with high levels of need for achievement are more apt to come from the middle class than from either the lower or upper classes, have better memory for incompleting tasks, are more apt to volunteer as subjects for psychological experiments, are more active in college and community activities, choose experts over friends as working partners, are more resistant to social pressure and cannot give accurate reports of what their "inner concern" with achievement is (McClelland, Atkinson, Clark, & Lowell, 1953 & Atkinson, 1958).

Finally, in a large cultural study, McClelland has found a strong relationship between high levels of need for achievement and an interest in enterprise which requires moderate or calculated risks, rather than safe or highly speculative undertakings (McClelland, 1961).

The implications of the theory of achievement motivation for the understanding of consumer behavior are great. Many aspects of purchase and consumption behavior are carried out in a manner that should allow an individual's level of need for achievement tendency to operate as a determinate of behavior. For example, it can be anticipated that there will be differential behavior in the selection of style and color of clothing, the degree of participation in a clothing fad, brand purchase behavior, and willingness to associate with various store images.

For the consumer in the market place, it is reasonable to suggest that not all purchase situations will be conducive to allowing n Achievement to operate as a behavioral determinate. This is due to the belief that when (P_s) and (I_s) are at relatively low levels and hence, when combined multiplicatively with (M_s), result in low levels of the strength of the motive to achieve (T_s). However, at least two situations would seem to effect the magnitude of (P_s) and (I_s) and hence, allow n Achievement (M_s) to operate as a behavioral determinate. These two situations (of which there may be many more) are situations where reference group influence is operating and situations where there are elements of social and

physical risk. These two situations would seem to allow an Achievement (M_s) to operate as a behavioral determinate due to the fact that the individual should perceive himself responsible for the outcome and there is some degree of risk concerning the possibility of success.

These two situations served as the basis for the questions to be explored in this study and the methodology.

The following questions are concerned with exploring the relationship between achievement motivation and the assortments of goods and services held and desired by consumers.

1. Do high need achievers, because of their preference for situations of moderate risk, prefer to shop for clothing in specialty stores and department stores as opposed to chain and "discount" stores?
2. Do high need achievers, because of high risk avoidance, tend to avoid products that would identify them with small, venturesome segments of the population?
3. Are high need achievers more apt to smoke cigarettes than people with lower levels of need for achievement as reported in earlier research?
4. In general, are brand names more important to high need achievers than for people with lower levels of need for achievement?

Research Design

The sample for this study consisted of 100 male juniors and seniors enrolled in an introductory marketing course at the University of Illinois, Urbana. Females were excluded because of repeated and consistent reports of instability of standard measures for assessing need for achievement levels for females (Atkinson, 1958). The data from thirteen subjects were incomplete leaving the actual sample of 87 subjects.

The first step in this study was the administration of a Consumer Preference Questionnaire. This questionnaire was developed especially for this study and was designed to reveal qualitative distinctions in the goods and services owned and desired by respondents. Questions were directed towards identifying types of goods owned, where purchased, and types of goods and services desired. For a question to be included in the questionnaire, it had to be related to a situation where reference group influence could operate or where varying levels of perceived risk could operate. The questionnaire was validated by extensive post-questionnaire interviews with individual respondents to ascertain that answers to individual questions were an accurate reflection of actual behavior. In its final form, the Consumer Preference Questionnaire was administered in a classroom situation.

The second step was to expose the same subjects used in the first step to a set of four TAT stimuli. The same four pictures have been used in many of McClelland's validation studies. This is also the most widely used set of stimuli in studies of college groups. The procedure for administration was identical to that recommended by Atkinson (Atkinson, 1958). The second step was also administered in a classroom situation approximately two weeks after the administration of step one.

The protocols elicited by the use of the TAT stimuli were scored for achievement motivation using the manual developed by McClelland and others (Atkinson, 1958). The scoring judge was trained by using the self-administered training procedure developed by Smith and Feld (Smith & Feld, 1958). Attempts to find another trained judge failed. However, the same judge had a scoring reliability of .85 when approximately ten percent of the protocols were rescored one month later.

The Consumer Preference Questionnaire was scored for individual responses. These responses were then paired with the same individual's level of need for achievement as measured in step two.

Results

In analyzing the results of this study, consistent patterns of behavior were looked for rather than statistical differences between groups for one product or situation. As more is learned about the need for achievement as it relates to goods and services, studies with very specific hypotheses lending themselves to statistical analysis will be more useful.

For ease of interpretation, the data is reported in nine categories. Each category is designed to give understanding to the influence of varying levels of need for achievement on purchase behavior. Only data for subjects with the highest and lowest levels of need for achievement is used in the following comparisons. If any differences do exist, this procedure should allow the differences to emerge. From the total distribution of need for achievement scores, the top seventeen percent are compared with the lowest eighteen percent. These percents were chosen because of natural breaks in scores on n Achievement. The middle sixty-five percent of distribution is not used in the comparisons unless specifically noted.

Ownership of a Broad Range of Products

Ownership was established by asking if the product was already owned. If the product was not presently a part of the person's assortment of goods, he was asked to estimate how likely it was that he would own such a good if he could afford it. He could respond, "Very Likely," "Maybe," "Not at all Likely," or "Undecided."

No noticeable differences in ownership or anticipated ownership appeared for the following goods and services:

- Color television
- Model train set
- Scotch whiskey
- Subscription to the Wall Street Journal
- The purchase of a new long play popular record
- Services of an interior decorator
- Bridge lessons
- Electric carving knife
- Flower garden book

TABLE I

Ownership and Purchase Intentions for
High and Low Need Achievers

	Very Likely		Maybe		Not at all Likely				Undecided		Already Own	
	High n-ach	Low n-ach	High n-ach	Low n-ach	High n-ach	Low n-ach	High n-ach	Low n-ach	High n-ach	Low n-ach	High n-ach	Low n-ach
Scuba Diving Outfit	3	1	4	2	6	13	2	0	0	0	0	0
Motor Bike	4	2	4	2	6	11	0	0	0	1	1	1
Water Skis	3	0	7	5	3	8	0	0	0	2	3	3
Snow Skis	5	1	6	6	4	7	0	0	0	0	2	2
Camping Equipment	7	0	5	6	2	9	0	0	0	1	1	1
"Mod" Clothing	0	0	2	7	12	8	1	1	1	0	0	0
Home Workshop	6	2	6	4	3	10	0	0	0	0	0	0

However, for a category of goods referred to as "outdoor sports," noticeable differences appear. Relationships seem to be suggested between the products listed in Table 1 and levels of need for achievement.

Except for the last two products, the products can be classified as "outdoor sports." For these five "outdoor sports" products, there is a tendency for the person with High n Achievement to express higher levels of purchase intentions. The same is also true for the product category - "camping equipment."

The reverse is found for the category - "mod clothing." Subjects with High n Achievement expressed considerably less likelihood for the purchase of such clothing.

Clothing Brands and Place of Purchase

This category of information was obtained by asking questions designed to identify the last brand of clothing purchased and shopping patterns for clothing.

Subjects with High n Achievement expressed only a very slight difference from subjects with Low n Achievement in their dress shirt preferences. Van Heusen dress shirts were slightly more preferred by subjects with High n Achievement. No noticeable differences were observed for purchases of Arrow, Manhattan, J. C. Penney, Gant and Hathaway shirts.

There is a noticeable difference, however, between purchases of men's suits, depending on the level of n Achievement. The data reported in Table II indicate that subjects with High n Achievement have a clear preference for quality men's store brands which far exceeds that of subjects with low levels of need for achievement. There does not appear to be any indication of a clear preference for subjects with Low n Achievement.

There is a similar, but less clear-cut tendency in the reported purchase of the last sport coat. The data reported in Table II indicates that, while differences between subjects with High and Low n Achievement exist, these differences are not as pronounced as those differences reported for suit purchases.

When asked to describe how much money is spent on a suit, in relation to resources, no differences in behavior were noted. However, there was a slight tendency for subjects with High n Achievement to indicate that they preferred to find a store they like and to go to it for all clothing accessory purchases, as opposed to subjects with Low n Achievement who expressed a higher tendency to compare between several stores for each purchase. No differences were noted in the type of service preferred when shopping for clothing, both groups preferring to have assistance available if needed. None of the subjects with Low n Achievement stated that their present best clothing was "old, but adequate." However, three out of fifteen subjects with High n Achievement did describe their best clothing this way.

TABLE II

Place of Purchase of Last
Suit and Sport Coat

	Suit		Sport Coat	
	High n-ach	Low n-ach	High n-ach	Low n-ach
Hart-Schaffner & Marx	0	4	NA	NA
Botany 500	1	2	NA	NA
A quality men's store brand	13	3	11	7
A low-prices men's store brand	1	3	1	2
J. C. Penney, or Sear, Roebuck	0	1	2	2
Department Store Brand	0	0	0	0
Robert Hall	0	0	0	0
Other	0	1	0	1

NA = not applicable

Toothpaste Preferences

Toothpaste preferences were obtained by asking subjects which brand of toothpaste they purchased most often. Responses to this question are presented in Table III.

The data reported in Table III suggest that there is a very slight relative preference for Crest toothpaste for subjects with Low n Achievement, and a very slight relative preference for McLeans toothpaste for subjects with High n Achievement.

TABLE III

Toothpaste Brand Preferences

	High n-ach	Low n-ach
Colgate	3	1
Ultra-Bright	2	3
Crest	4	9
McLeans	5	2
Other	1	1

Cola and Beer Preferences

By asking what brand is preferred when drinking cola and drinking beer, no noticeable differences between subjects with High and Low Achievement were noted. Their preferences were both distributed in approximately the same proportions.

Willingness to Try New Products

Subjects were asked to check a statement that best described their willingness to try entirely new products and then that statement which best described their willingness to try a new brand of an established product. Subjects with High n Achievement were only slightly more willing to try entirely new products and no differences were noted for willingness to try a new brand of an established product between subjects with High and Low n Achievement.

Myself As A Customer

A picture of the subject as a shopper was constructed by having subjects select descriptions of themselves in several different contexts. Nine out of fifteen subjects with High n Achievement described themselves as "rational," while only four out of fifteen with Low n Achievement described themselves as "rational." However, nine out of fifteen subjects with Low n Achievement described themselves as "conservative," while only three out of fifteen subjects with High n Achievement described themselves as "conservative."

One-half of the subjects with Low n Achievement believed that "you get exactly what you pay for," while only one subject with High n Achievement believed that statement. However, fourteen out of the fifteen subjects with High n Achievement believed "the relationship between price and quality is often misleading." Only one-half of the subjects with Low n Achievement believed that statement.

Given a choice of purchasing a RCA, 19 inch, black and white, portable television set from a well-known department store at \$139.95 or the same set from the local "discount" store at \$126.37, subjects with Low n Achievement expressed a very slight preference for making the purchase from the department store at \$139.95.

Importance of Brand Names

The importance of brand names was ascertained by first asking subjects to choose a statement that indicated their feelings about the importance of brand names and then to indicate whether a brand name is important in the purchase of eighteen products.

Seven of the sixteen subjects with Low n Achievement felt that "brand names were very important to insure that they were buying the right product." Only three of the fifteen subjects with High n Achievement felt brand names were that important.

Out of the eighteen products that were especially chosen with respect to the potential usefulness of brand names in purchase decision, only four

products show even the slightest difference between subjects with High and Low n Achievement. Three of the sixteen subjects with Low n Achievement felt that brand name was important in the purchase of aspirin. However, no subject with High n Achievement felt that brand name was important in the purchase of aspirin.

In the purchase of casual slacks, nine of the sixteen subjects with Low n Achievement felt that brand name was important, while four of the fifteen subjects with High n Achievement felt that way.

Limited differences are seen also in paint for a room and a coffee pot. For both products, two of the sixteen subjects with Low n Achievement felt that brand name was important. However, we find contrary to slacks and aspirin, six of the fifteen subjects with High n Achievement feeling that brand name was important in the purchase of paint for a room and a coffee pot.

Automobile Preferences

When asked to rank order preferences for different types of automobiles, some differences were evident. The values in Table IV represent the mean rank value for each automobile.

Subjects with High n Achievement differed from subjects with Low n Achievement on automobile preferences in two instances. Luxury automobiles were ranked 7th by subjects with High n Achievement but 3rd by subjects with Low n Achievement. A somewhat less dramatic difference is noted for Intermediate size automobiles with subjects with High n Achievement ranking them 2nd while subjects with Low n Achievement ranking them 5th.

Smoking Behavior

The smoking behavior of subjects with the very highest n Achievement and the very lowest n Achievement is almost identical. About one-third of each group reported that they regularly smoked cigarettes. However, when comparing two groups obtained by drawing a line at the median need achievement score, a different result is obtained. Forty percent (17/43) of subjects in the top fifty percent of n Achievement distribution reported that they regularly smoked cigarettes while only 16 percent (7/37) of subjects in the bottom fifty percent reported the regular smoking of cigarettes. The distribution of smokers is somewhat skewed with the greatest concentration falling between the upper 20 and 50 percent limits of the need for achievement distribution. The same type of distribution was also found for pipe smoking, although on a much more reduced scale.

Discussion

The results of this study suggest that there is a relationship between levels of need for achievement and consumer behavior. This difference shows up as generally positive and consistent answers to three of the four questions posed earlier.

TABLE IV

Automobile Rank Order Preferences

(High n Achievement)

Mean Rank Value*	Automobile Types
2.2	Small domestic sports (Mustang, Camero, Javlin, etc.)
3.2	Intermediate (Chevelle, Fairlane, Belveder, etc.)
3.2	Full Size - Big 3 (Chevrolet, Ford, Plymouth)
3.9	Imports (all)
4.1	Full Size (Buick, Chrysler, Mercury, Oldsmobile, Pontiac)
5.8	Compact (Valiant, Chevy II, etc.)
5.9	Luxury (Cadillac, Lincoln, Imperial, Thunderbird)

Automobile Rank Order Preferences

(Low n Achievement)

Mean Rank Value	Automobile Types
2.6	Small domestic sports (Mustang, Camero, Javlin, etc.)
3.3	Full Size (Buick, Chrysler, Mercury, Oldsmobile, Pontiac)
3.5	Luxury (Cadillac, Lincoln, Imperial, Thunderbird)
4.1	Full Size - Big 3 (Chevrolet, Ford, Plymouth)
4.3	Intermediate (Chevelle, Fairlane, Belveder, etc.)
4.5	Imports (all)
6.5	Compact (Valiant, Chevy II, etc.)

*1 = most preferred

7 = least preferred

The first question dealt with the preference for shopping in specialty and department stores for clothing for subjects with High n Achievement. There is a consistent and strong pattern to indicate that subjects with High n Achievement have a preference for purchasing suits and sport coats from specialty clothing stores. No such clear-cut preference was noted for the purchase of dress shirts. It appears that suits and sport coats purchased from a "quality" men's store represent a moderate amount of risk for subjects High n Achievement. This may be a localized phenomenon but is certainly consistent with the previous research of Atkinson reported earlier in this paper.

The second question suggested that subjects with High n Achievement would avoid products that identified them with small, venturesome segments of the population. The main question dealing with this issue was the ownership of "Mod" clothing. However, possibly because of high levels of perceived risk, subjects with Low n Achievement were much more apt to own or intend to purchase such clothing.

Closely related to this question is the finding that subjects with High n Achievement expressed a preference for ownership of goods that can be characterized as "active outdoor sports or recreation." Based on previous research, it appears that active outdoor sports are perceived as being of moderate risk. Subjects with Low n Achievement may perceive low probability of success and low incentive to succeed in such activities, and hence, be more apt to avoid them.

That subjects with High n Achievement almost uniformly put the lowest possible ranking on luxury cars is hard to explain. They may be seen as a high risk item or that luxury cars are not needed to prove something.

The third question suggested that subjects with High n Achievement would be more apt to smoke cigarettes. It is difficult to explain why the second quartile (subjects with moderately High n Achievement) should contain such a high number of cigarette smokers. While the evidence of this study supports the earlier work of Carney (Carney, 1967), the explanation must lie in the interaction between moderately High n Achievement and the perceived risk of smoking.

There is little evidence to support the suggestions of the fourth question that brand names would be more important to subjects with High n Achievement. No difference was noted between subjects with either High or Low n Achievement. If subjects with High n Achievement are more rational, this alone should cause more or less reliance on brand names. It appears, however, that brand names do not influence the amount of perceived risk and/or are not subject to reference group influence. Hence, n Achievement does not enter in as a behavioral determinate in this situation.

Related to the above questions is the finding that subjects with High n Achievement consider themselves as "rational" shoppers. In addition, subjects with High n Achievement appear to prefer to find a single store for all purchases rather than shopping around for each purchase. One of the more interesting findings of this study was the preference of subjects with Low n Achievement for purchasing a television set from a well-known department store, rather than for less money at a discount store.

Relating the evidence of this study to the formulation of Atkinson (Atkinson, 1964), it appears that, for many of the products and situations, (P_s) and (I_s) are at low levels. Hence, when combined with (M_s), no matter what its magnitude, results in finding no noticeable differences between groups with High and Low n Achievement. However, the combination of (P_s) and (I_s) do appear to be at moderate levels for "active outdoor sports" products. In addition, it appears as if certain clothing stores are selected and certain shopping patterns followed because the combination of (P_s) and (I_s) associated with these stores and shopping patterns is moderate. Therefore, in these two later cases, when combined with High levels of (M_s), a differential tendency to approach success is noted.

Implications

While not explaining a large portion of consumer behavior, the theory of achievement motivation does appear to have the potential for accounting for behavior in some situations. With the limited amount of evidence reported here, it would be appropriate for more studies to be conducted so that stores and manufacturers may better tailor their appeals, stores, and products to a specific group of customers.

Based on this study, the main group of stores who could profit from recognizing the existence of a group of customers with High n Achievement are men's clothing and the clothing departments of department stores. Based on what we know about people with High n Achievement, people with High n Achievement may be more apt to patronize an institution that uses appeals of excellence, just exciting enough to generate moderate amounts of perceived risk, liberal use of positive achievement words, and appropriate shades of color in stores and merchandise.

The main group of manufacturers who could profit are those making the active outdoor sports equipment. They are most likely to satisfy this group by generating just the right level of perceived risk, neither too high or too low. There is enough evidence in this study to suggest that a very large portion of people engaged in outdoor, active sports may be high in need for achievement. If so, this could seriously alter the promotion and distribution of these products.

While this study should be replicated on other groups and additional studies should be conducted on a variety of relationships between n Achievement and consumer behavior, it is unlikely that this area of interest will be heavily researched. The reason for this lack of research is the difficulty of administering and scoring projective tests. Unfortunately, there is no objective measure at present that has an acceptable correlation with the validated method reported and used in this study.

References

- Atkinson, J. W. An Introduction to Motivation. New York: American Book-Van Nostrand-Reinhold, 1964.
- Atkinson, John W. (ed.), Motives in Fantasy, Action and Society. Princeton: D. Van Nostrand Company, Inc., 1958).
- Atkinson, J. W. & Feather, N. T. (eds.), A Theory of Achievement Motivation. New York: John Wiley & Sons, 1966.
- Atkinson, John W. Motivational Determinants of Risk-Taking Behavior. Psychological Review, 1957, 64, 359-372.
- Carney, Richard E. Sex Chromatin, Body Masculinity, Achievement Motivation & Smoking Behavior. Psychological Reports, 1967, 20 (3, pt. 1), 859-866.
- Costello, C. G. Two Scales to Measure Achievement Motivation. Journal of Psychology, 1967, 66 (2), 231-235.
- Knapp, Robert H. n Achievement and Aesthetic Preference. In Atkinson, John W. (ed.), Motives in Fantasy, Action and Society. Princeton: D. Van Nostrand Company, Inc., 1958, 367-372.
- Littig, L. W. The Effect of Motivation on Probability Preferences and Subjective Probability. Unpublished doctoral thesis, 1959, University of Michigan.
- Litwin, G. H. Motives and Expectancies As Determinants of Preference for Degrees of Risk. Unpublished honors thesis, 1958, University of Michigan.
- McClelland, D. C. The Achieving Society. Princeton: D. Van Nostrand Company, Inc., 1961.
- McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. The Achievement Motive. New York: Appleton-Century-Crofts, 1953.
- McClelland, D. C., Clark, R. A., Roby, T., & Atkinson, J. W. The Projective Expression of Needs: IV. The Effect of the Need for Achievement on Thematic Apperception. Journal of Experimental Psychology, 1949, 39, 242-251.
- McClelland, D. C. & Liberman, A. M. The Effect of Need for Achievement on Recognition of Need-related Words. Journal of Personality, 1949, 18, 236-251.
- Smith, Charles P. & Feld, Sheila. How to Learn the Method of Content Analysis for n Achievement, n Alliliation and n Power. In Atkinson, J. W. (ed.), Motives in Fantasy, Action and Society. Princeton: D. Van Nostrand Company, Inc., 1958, Appendix I.

SOCIOECONOMIC AND PERSONALITY CHARACTERISTICS
AS THEY RELATE TO
ECOLOGICALLY-CONSTRUCTIVE PURCHASING BEHAVIOR

Thomas C. Kinnear
University of Western Ontario
James R. Taylor
University of Michigan
and Sadrudin A. Ahmed¹
Bishops University

Introduction

Personal consumption has important ecological implications. Individuals in North America create as much solid waste as manufacturing does (Lynch & Chandler, 1971). Automobile exhaust accounts for a significant proportion of air pollution. Phosphate in laundry detergents has been identified as a significant source of water pollution in the Great Lakes (International Joint Committee, 1971). Glass bottles and aluminum cans present special disposal problems because they do not deteriorate under normal disposal methods. Pesticides and fertilizers for home use also contribute to the pollution of water resources. So it appears that wisdom buyers exercise in their purchase decisions with regard to these types of products can either help maintain the environment or contribute to its deterioration.

The purpose of this paper is to analyze the relationship between the socioeconomic and personality characteristics of respondents and the amount, if any, of ecological concern in their buying behavior. Examination of the relationship of socioeconomic and personality characteristics to consumer purchase behavior has received much attention in the literature. In 1968 Frank provided a discussion of the relative levels of success in using these types of predictors (Frank, 1968). At that time he concluded that socioeconomic and personality characteristics had not proven to be very useful predictors. Since that time some success has been obtained in relating personality characteristics to product choices. Fry identified relationships of significant magnitude between personality variables and cigarette brand choice (Fry, 1971). Alpert also concluded that personality is a useful construct in identifying the determinants of product choice (Alpert, 1972).

Both socioeconomic and personality factors can be conceived of as important when brands of products with ecological implications are chosen. Purchasing of ecologically safe products may be the norm of a particular age group, may require higher incomes to cover the extra costs or may require a person to have more education so he can understand and appreciate the importance of his individual consumption patterns. Further, one's purchasing pattern in relation to pollution may be a way for consumers to display their personality structure.

Here we are concerned with the ability of certain predictors to aid in classifying respondents into groups of doers and nondoers according to whether or not they choose products which are ecologically constructive. The Automatic Interaction Detector (AID) and Multiple Classification Analysis (MCA) programs are the primary analysis techniques utilized for this purpose. Three behavioral measures are analyzed individually. These measures are:

- (1) The use of a nonphosphate laundry product
- (2) The use of returnable bottles

- (3) The existence of an unusual shopping pattern in order to purchase nonpolluting products, i.e., did the consumer alter her usual shopping procedure with the intent of obtaining nonpolluting products?

Each of these measures was coded "1" if it occurred and "0" if it did not occur. Since there is only one interval on the dependent variable, the metric assumption of both AID and MCA is satisfied.

The Data

The behavioral data utilized in this paper was collected by means of a questionnaire that was mailed to 698 members of the Canadian Family Opinion--University of Western Ontario Consumer Panel. Useable questionnaires were returned by 500 panel members, constituting 72 per cent of the panel. Comparison of the socioeconomic characteristics of panel members who answered the questionnaire with those who did not answer it did not indicate any significant differences.

The personality measures were collected previously by Sadrudin A. Ahmed as part of his dissertation research. He obtained data for the personality scales from only 348 of the 500 respondents. The others are treated as missing data.

Description of the Independent Variables

Twenty independent variables were available as possible predictors in separating the doers and nondoers. Seven of the predictors were socioeconomic. They are:

- (1) Age of Wife
- (2) Presence of Children
- (3) Education of Wife
- (4) Education of Husband
- (5) Employment of Wife
- (6) Occupation of Principal Wage Earner
- (7) Family Income

Twelve of the predictors were provided by scores on standard personality scales. These scales are listed below and described in the following section:

- (8) Aggression
- (9) Anxiety
- (10) Depression
- (11) Desirability
- (12) Dominance
- (13) Harm Avoidance
- (14) Play
- (15) Rebelliousness
- (16) Self-esteem
- (17) Sentience
- (18) Understanding
- (19) Tolerance
- (20) The final predictor--Perceived Consumer Effectiveness--is a measure of the extent to which a respondent believes that an individual

consumer can be effective in pollution abatement. This variable was obtained from responses to the following statement in the questionnaire: "It is futile for the individual consumer to try to do anything about pollution."

Those who strongly disagreed with this statement were assigned to the highest category of Perceived Consumer Effectiveness; those who strongly agreed were assigned to the lowest category; and respondents were assigned to middle categories.

THE STANDARD PERSONALITY SCALES

The measurement of personality in this paper follows the trait-type tradition of Murray (Murray, 1938). The basic assumptions of trait theory are that personality is made up of certain definite attributes that are common to many people, vary in amount, and can be measured by indicators. Further, it is assumed that traits are relatively stable and exert generalized causal effects on behavior (Mischel, 1968).

The twelve personality scales utilized were developed by Douglas N. Jackson to discriminate between individuals in the normal range of psychological makeup. The Aggression, Desirability, Dominance, Harm Avoidance, Play, Sentience, and Understanding scales are taken from his Personality Research Form (Jackson, 1967). Self-esteem, Tolerance and Anxiety scales are derived from the Jackson Personality Inventory (Jackson, 1970), and the Rebelliousness and Depression scales come from Jackson's Differential Personality Inventory (Jackson, 1970). A description of the high scorer on each of these scales is presented in the Appendix. All of the Jackson personality scales contained from fifteen to twelve categories. Since the many categories would cause degrees of freedom problems in MCA runs, all personality scales were condensed into five categories plus a missing data category.

The twenty independent variables were analyzed using the AID procedure for each of the three behavioral measures under consideration. Tables 1, 2, and 3 show the results of the AID analysis on the use of a nonphosphate brand of soap, the use of returnable bottles, and the existence of an unusual shopping pattern, respectively. Each predictor was treated as nominal in order to allow for the possibility of curvilinear splits. These tables were examined to determine the best predictors for each dependent variable. Table 4 presents the predictors that were selected to be run in the MCA analysis for each dependent variable.

Results

The estimates produced by MCA are shown in Tables 5, 6, and 7. A number of calculations are presented for each predictor. The unadjusted coefficient is the deviation from the grand mean associated with the categories of any predictor taken by itself. The adjusted coefficient is the deviation from the grand mean associated with the categories of any predictor taking into account the effects of all other predictors. Because doers are coded "1" on the dependent variable and nondoers are coded "0", a special interpretation of the MCA results is possible. The grand mean can be interpreted as the probability that a respondent selected at random from the sample will undertake the behavior being examined in the MCA run. The coefficients produced by MCA then represent changes in this overall probability associated with the knowledge that a respondent falls into a particular category of a predictor variable.

Use of a Nonphosphate Laundry Product

The overall probability that a respondent will use a nonphosphate brand is .367. Table 5 shows the adjusted effects of each predictor. The most important predictors are:

PERCEIVED CONSUMER EFFECTIVENESS

Those scoring in the highest category of Perceived Consumer Effectiveness are substantially more likely to purchase a nonphosphate laundry product than those in the lower categories. The adjusted probability of this purchase increases by .164 to .531 with knowledge that a respondent is in the highest category. Those scoring in the middle three categories in Perceived Consumer Effectiveness have a probability of purchase much below the overall average. Those in the lowest category are above average.

PRESENCE OF CHILDREN

Those respondents with children in all three age groups (under 6, 6-12, and 13-18) and those with children under 6 and ages 13-18 show an adjusted probability of .683 and .631 that they will purchase a nonphosphate soap. The latter category has a very small sample size associated with it, and this fact may have affected the coefficient. There appears to be something in the dynamics of families of these types that is associated with the purchase of a nonphosphate brand. The discussion section of this paper examines these dynamics in more detail.

INCOME

Only the highest income category has a meaningful effect on the probability of this purchase. Those with incomes of \$15,000 and over have an adjusted probability of purchase of .469, which is .102 above average.

HARM AVOIDANCE

An increasing Harm Avoidance score is associated with a relatively stable pattern in the probability of this purchase until the highest category is reached. At this point the probability falls substantially below average to .199.

All other predictors yield small or relatively random changes in probability.

Use of Returnable Bottles

Table 6 shows the MCA results associated with the use of returnable bottles. The overall probability that a respondent will use returnable bottles is .589. The most important predictors are:

PERCEIVED CONSUMER EFFECTIVENESS

Those scoring highest in Perceived Consumer Effectiveness have an adjusted probability of purchase of returnable bottles of .672, which is .083 above the average. Those scoring the lowest have a probability of .075 below the average. Those in the middle three categories are also below average. There is no evidence of a curvilinear relationship in this instance.

PRESENCE OF CHILDREN

Those with children in all three age groups are substantially more likely to purchase returnable bottles than those without children. The probability of this purchase increases by .175 to .764. Those families with children under 19 are associated with a decrease in probability of .104 to .485. The same dynamics, as noted previously, seem to be operative again (see Discussion section). All other predictors yield small or relatively random changes in probability.

Existence of an Unusual Shopping Pattern

Table 7 shows the MCA results associated with the existence of an unusual shopping pattern. The overall probability that a respondent will indicate the existence of such a pattern is .150. The most important predictors are:

PERCEIVED CONSUMER EFFECTIVENESS

Those scoring in the highest category of Perceived Consumer Effectiveness are again more likely than the average to undertake an unusual shopping pattern. The highest scorers have an adjusted coefficient of .064, which yields a probability of .214. Those scoring in the middle three categories have below average probability of undertaking an unusual shopping pattern. The nineteen respondents in the lowest category are again above average. The possible curvilinear pattern that was noted for the use of a nonphosphate laundry product is also evident here.

AGGRESSION

Low scorers in the Aggression category are much less likely than the average to indicate an unusual shopping pattern. The lowest category has an adjusted probability of .040. Those scoring highest have an adjusted probability of .234. It appears that as aggression increases, the respondents are more likely to undertake unusual shopping behavior.

TOLERANCE

The highest scorers in the Tolerance category are much more likely than the average to indicate an unusual shopping pattern. Their adjusted probability is .132 above average, yielding a probability of .282.

INCOME

Those with incomes of \$15,000 and over have an adjusted probability of purchase of .015. This is .135 below the average. No other income category produces marked deviations from average probability.

EDUCATION OF HUSBAND

Those respondents whose husbands have a university or college education are more likely than the average to indicate an unusual shopping pattern. Those whose husbands have only an elementary school education or those who have no husband are less likely than the average to indicate an unusual shopping pattern.

Discussion

Role of Perceived Consumer Effectiveness

The role of Perceived Consumer Effectiveness is very similar for all three behavioral measures. Respondents scoring in the highest category of Perceived Consumer Effectiveness have a higher probability than the average of undertaking ecologically constructive behavior. Those in the middle three categories are associated with a below-average probability. The lowest category is associated with an above-average probability that they will use a nonphosphate laundry product and will go out of their way to buy nonpolluting products. These latter results are associated with a small sample size. If they are, indeed, valid, then the relationship between Perceived Consumer Effectiveness and these measures of ecological concern is curvilinear. There does not appear to be an intuitively appealing explanation of this possible curvilinear relationship.

If the latter results are not valid, then an intuitively appealing explanation is available. That is, as respondents perceive that consumers are increasingly effective in pollution abatement, then they are more likely to demonstrate concern for ecology.

Presence of Children

As we have seen, respondents with children in all three age group categories (under 6, 6-12, and 13-18) indicate a much higher than average probability that they will use a nonphosphate laundry product and returnable bottles than do respondents without children. There appears to be something in the dynamics of families of this type that is associated with their buying activities, although these dynamics are not clear. It may be that with the increase in the number of children in a family, the possibility that one of the children will successfully influence the mother to undertake ecologically-maintaining purchasing behavior increases. On the other hand, it may be that interaction among the children gives rise to a unified position that the mother should purchase in an ecologically-constructive fashion. This unified pressure may be stronger than the pressure exerted by children in other types of families. The data collected for this paper do not bring any evidence to bear on the possible dynamics, and therefore, no definitive statement can be made.

Use of a Nonphosphate Laundry Product

In addition to Perceived Consumer Effectiveness and the presence of children, we have found that the use of a nonphosphate laundry product is related mainly to the presence of an income of \$15,000 and over. High income may play a facilitating role, i.e., respondents in this income range may be less concerned with the possible extra costs of nonphosphate laundry products.

The deviations associated with Harm Avoidance present an interesting pattern. The lowest scorers on this trait have a small negative deviation. As Harm Avoidance scores rise, the deviations become increasingly positive until the highest Harm Avoidance category is reached. At this point the deviation becomes substantially negative.

The increase in probability of purchase through four categories is an intuitively appealing result. It is expected that a person concerned about being harmed by pollution would demonstrate more concern about the ecology. However, those highest in Harm Avoidance are least concerned with ecology. It appears that when Harm Avoidance related to pollution becomes extremely high, a person reacts by ignoring it to a great extent. They think that the solution to potential pollution harm is to avoid thinking or doing much about it. This finding is consistent with research findings on the use of fear appeals in communications. In these studies it has been found that strong appeals to fear are less effective in persuasion than minimal appeals because too much tension is created by the strong appeals (Jackson, 1967). People exposed to a strong fear appeal tune out the communication. In much the same way, those scoring highest in the personality trait, Harm Avoidance, tune out potential pollution concern.

Thus, ability to predict whether a respondent will use a nonphosphate laundry product is most enhanced by knowledge of her Perceived Consumer Effectiveness score, the presence of children in her family, and her family income. For example, a respondent who is in the highest category of Perceived Consumer Effectiveness, has children in all three age groups, has an income of \$15,000 or more, and is in the middle category of all other predictors would have a .944 probability of purchasing a nonphosphate laundry product. The calculation is as follows:

Estimated probability = Grand mean

- + effect of being highest category of Perceived Consumer Effectiveness
- + effect of having children in all three age groups
- + effect of being in the middle category of Aggression, Self-esteem, Tolerance, Understanding, Dominance, Harm Avoidance, and Play.

Therefore:

$$\begin{aligned}
 \text{Estimated probability} &= .367 + .164 + .316 + .102 \\
 &\quad + (-.059 + .069 - .049 + .058 \\
 &\quad \quad - .035 + .032 - .021) \\
 &= .367 + .164 + .316 - .102 - .005 \\
 &= .949 - .005 \\
 &= .944
 \end{aligned}$$

Use of Returnable Bottles

The knowledge of Perceived Consumer Effectiveness scores and the presence of children category have a marked effect on the ability to predict the use of returnable bottles. For example, a respondent who is in the highest category of Perceived Consumer Effectiveness, has children in all three age groups, and is in the middle category of all other predictors would have a .705 probability of using returnable bottles. The calculations are:

Estimated probability = Grand mean

- + effect of being in the highest category of Perceived Consumer Effectiveness
- + effect of having children in all three age groups
- + effect of being in the middle category of Income, Aggression, Depression, Tolerance, Understanding, Anxiety, Dominance, and Play

Therefore:

$$\begin{aligned}
 \text{Estimated probability} &= .589 + .083 + .173 \\
 &+ (-.026 - .089 + .061 + .015 \\
 &\quad - .053 - .038 - .084 + .079) \\
 &= .589 + .083 + .173 - .140 \\
 &= .845 - .140 \\
 &= .705
 \end{aligned}$$

Existence of an Unusual Shopping Pattern

Besides the Perceived Consumer Effectiveness score, Aggression, Tolerance, Income, and Education of husband are the most useful predictors of the existence of an unusual shopping pattern. Aggression was not markedly related to any other behavioral measure. It appears that the alternation of one's shopping pattern is a very different type of decision than that for the other two behavioral measures. Respondents who are highly assertive are more likely to undertake this behavior.

High scorers in Tolerance are more likely to alter their shopping pattern. It appears that the alteration of one's habitual shopping pattern requires an above average amount of openness to new ideas and ways.

It was noted previously, that respondents with incomes of \$15,000 or over were more likely to purchase a nonphosphate laundry product and were more likely to have higher scores on the ecological concern index. It is interesting to note that these respondents are much less likely than the average to undertake an unusual shopping pattern. Apparently, these respondents are willing to purchase ecologically-constructive laundry products but are unwilling to go out of their way to do so. People with high incomes are, perhaps, too used to convenience in their shopping to expend a special effort.

Respondents whose husbands have a college or university education are more likely to undertake an unusual shopping pattern. This result, in combination with the results noted above about high income families, is confusing. It is logical to expect that college educated people would have higher incomes. Why then would a college education be positively associated with an unusual shopping pattern while high incomes are negatively associated with it? This pattern of results holds also when the education and income levels are examined before adjustments are made for other factors (see Table 7). No clear answer is available.

All four variables--Perceived Consumer Effectiveness, Aggression, Tolerance, and Education--influence the probability of undertaking an unusual shopping pattern. For example, a respondent who is in the highest category of Tolerance, whose husband has a college education, and is in the middle category of all other predictors, has a .486 probability of undertaking an unusual shopping pattern.

Estimated probability = Grand mean

- + effect of being in the highest category of Perceived Consumer Effectiveness
- + effect of being in the highest category of Aggression
- + effect of being in the highest category of Tolerance
- + effect of husband having a college education
- + effect of being in the middle category of Income, Age, Harm Avoidance, Play, and Rebelliousness.

Therefore:

$$\begin{aligned}
 \text{Estimated probability} &= .150 + .064 + .084 + .132 \\
 &\quad + .085 + (-.005 - .009 - .010 \\
 &\quad + .056 - .061) \\
 &= .150 + .064 + .084 + .132 \\
 &\quad + .085 - .029 \\
 &= .515 - .029 \\
 &= .486
 \end{aligned}$$

Footnotes

1. Thomas C. Kinnear is Assistant Professor Business Administration, The University of Western Ontario. James R. Taylor is Associate Professor of Marketing at The University of Michigan. Sadrudin A. Ahmed is Assistant Professor of Business Administration at Bishops University.

References

- Alpert, M.I. Personality and the Determinants of Product Choice. Journal of Marketing Research, 1972, 9, 89-92.
- Frank, R.E. Market Segmentation Research: Findings and Implications. In Frank Bass, Charles King, & Edgar Pessemier (Eds.), Applications of the Sciences in Marketing Management. New York: John Wiley and Sons, Inc., 1968.
- Fry, J.N. Personality Variables and Cigarette Brand Choice. Journal of Marketing Research, 1971, 8, 298-304.
- International Joint Commission. Report on the Pollution of the Great Lakes. Ottawa, Ontario: Queen's Printer, 1971.
- Jackson, D.N. Personality Research Forum. Goshen, N.Y.: Research Psychologists Press, Inc., 1967.
- Jackson, D.N. Differential Personality Inventory. London, Ontario: University of Western Ontario, 1970.

Jackson, D.N. Jackson Personality Inventory. London, Ontario, University of Western Ontario, 1970.

Lynch, P. & Chandler, R. National Environment Test. New York: Pocket Books, 1971, p. 109.

Mischel, L.W. Personality and Assessment. New York: John Wiley and Sons, Inc., 1968.

Murray, H.A. Explorations In Personality. Cambridge: Harvard University Press, 1938.

Appendix

Description of High Scorer on Personality

1. Aggression

Enjoys combat and argument; is easily annoyed; sometimes willing to hurt people to get his way; may seek to get even with people whom he perceives as having harmed him

2. Anxiety

Tense, restless, uneasy; tends to worry over inconsequential matters; more easily upset than is the average person; apprehensive about the future.

3. Depression

Is inclined to be downhearted and shows extreme despondence; considers himself to be inadequate; may be listless, remote, and preoccupied; looks at his future pessimistically

4. Desirability

Describes self in terms judged as desirability; consciously or unconsciously, accurately or inaccurately, present favorable picture of self in responses to personality questionnaire statements

5. Dominance

Attempts to control his environment and to influence or direct other people; expresses opinions forcefully; enjoys the role of leader and may assume it spontaneously

6. Harmavoidance

Does not enjoy exciting activities, especially if danger is involved; avoids risks of bodily harm; seeks to maximize personal safety

7. Play

Does many things "just for fun"; spends a good deal of time participating in games, sports, social activities, and other amusements; enjoys jokes and funny stories; maintains a light-hearted, easy-going attitude toward life

8. Rebelliousness

Will frequently be uncooperative, disobedient, and resistant when faced with rules and regulations; reacts against discipline and criticism

9. Self-esteem

Self-assured, egotistical, self-sufficient; confident in dealing with others; not easily embarrassed or influenced by others; imperturbable in interpersonal situations

10. Sentience

Notices smells, sounds, sights, tastes, and the way things feel; remembers these sensations and believes that they are an important part of life; is sensitive to many forms of experience; may maintain an essentially hedonistic or aesthetic view of life

11. Understanding

Wants to understand many areas of knowledge; enjoys logical thought; synthesis of ideas, verifiable generalizations; values thought, particularly when it is directed at satisfying intellectual curiosity or problem solving

12. Tolerance

Broad-minded, undogmatic, open-minded, accepts people even though their beliefs and customs may differ from his own; open to new ideas; free from prejudice

TABLE 1

USE OF NONPHOSPHATE LAUNDRY PRODUCT: PROPORTION

OF VARIATION EXPLAINABLE FOR EACH PREDICTOR

(Percentage)

Predictor	Group Number												
	1	2	3	4	5	6	7	8	9	10	11	12	13
Age of wife	.015	.016	.015	.006	.033	.005	.018	.053	.031	--	--	.041	.041
Presence of children	.022	.029*	.078*	.008	.002	.012	.035	.010	.016	--	--	.028	.028
Education of wife	.004	.001	.002	.005	.003	.000	.000	--	.004	--	--	.001	.001
Education of husband	.003	.001	.012	.023	.007	.010	.026	--	.013	--	--	.020	.020
Employment of wife	.010	.001	.045	.006	.007	.034	.073	.020	.001	--	--	.000	.000
Occupation	.010	.012	.020	.080*	.014	.047	.039	.060	.015	--	--	.015	.015
Income	.008	.008	.006	.052	.009	.002	.048	.034	.026	.026	.026	.025	.025
Perceived consumer effectiveness	.083*	.022	--	.002	.036	---	--	--	.012	.012	--	.012	.012
Aggression	.055	.099	.022	.037	.010	.016	.049	.090*	.017	.017	.045	.016	.016
Depression	.017	.011	.027	.047	.012	.031	.041	.046	.027	.027	.021	.010	.010
Self-esteem	.010	.017	.024	.046	.029	.045	.118	.072	.056*	.005	.005	.022	.022
Sentience	.008	.018	.018	.020	.025	.020	.086	.048	.021	.063	.063	.043	.043
Tolerance	.009	.019	.031	.020	.047	.046	.136*	.041	.030	.022	.022	.050	.050
Understanding	.010	.017	.010	.040	.060*	.011	.026	.048	.006	.020	.020	.016	.016
Anxiety	.015	.017	.039	.008	.029	.038	.040	.057	.006	--	--	.025	.025
Desirability	.010	.009	.020	.033	.013	.030	.103	.041	.015	.031	.031	.036	.036
Dominance	.006	.027	.018	.036	.058	.014	.025	.072	.038	.103*	.103*	.050*	.050*
Harmavoidance	.017	.028	.009	.040	.039	.013	.014	.043	.025	--	--	.043	.043
Play	.015	.019	.046	.029	.029	.050*	.015	.034	.032	--	--	.038	.038
Rebelliousness	.022	.024	.028	.011	.038	.042	.090	.065	.034	.034	.032	.037	.037

*AID split on this variable.

TABLE 2
 USE OF RETURNABLE BOTTLES; PROPORTION OF VARIATION
 EXPLAINABLE FOR EACH PREDICTOR
 (Percentage)

Predictor	Group Number											
	1	2	3	4	5	6	7	8	9	10	11	12
Age of wife	.014	.016	.001	.016	.047	.025			.016	.046		.013
Presence of children	.022*	.001	.002	.008	.060	.004			.006	.035		.002
Education of wife	.000	.017	.004	.004	.011	.007			.020	.004		.028
Education of husband	.000	.003	.003	.004	--	.011			.004	.013		.011
Employment of wife	.006	.001	.010	.017	--	.002			.016	---		.017
Occupation	.020	.054*	.013	.018	.054	.013			.014	.010		.010
Income	.009	.019	.012	.022	.007	.027			.024	.044		.029
Perceived consumer effectiveness	.013	.034	.023	.037	.020	.028			.030	.029		.042*
Aggression	.019	.013	.018	.019	.068	.027			.038	.009		--
Depression	.010	.016	.026	.047*	.008	.037			.025	.009		--
Self-esteem	.009	.006	.022	.018	.042	.022			.054	.025		--
Sentience	.014	.014	.027	.027	.040	.037			.050	.025		--
Tolerance	.006	.019	.009	.025	.020	.046			.068*	.012		--
Understanding	.021	.044	.018	.006	.034	.074*			.044	.011		--
Anxiety	.012	.008	.022	.016	.047	.042			.046	.047*		--
Desirability	.011	.025	.014	.016	.008	.059			.049	.044		--
Dominance	.018	.018	.028*	.002	.017	.028			.019	.014		--
Harmavoidance	.011	.008	.022	.022	.051	.029			.040	.029		--
Play	.015	.021	.017	.016	.088*	.043			.042	.021		--
Rebelliousness	.013	.011	.015	.019	.080	.034			.028	.018		--

*AID split on this variable.

TABLE 3

EXISTENCE OF AN UNUSUAL SHOPPING PATTERN: PROPORTION OF
 VARIATION EXPLAINABLE FOR EACH PREDICTOR

(Percentage)

Predictor	Group Number											
	1	2	3	4	5	6	7	8	9	10	11	
Age of wife	.006	.009	.019	.015	.036	.062		.006				.043*
Presence of children	.013	.018	.020	.021	.035	.040		.011				.033
Education of wife	.013	.015	.002	.001	.005	.003		.013				--
Education of husband	.020	.038*	.027	--	.013	.026		.006				.023
Employment of wife	.003	.005	.003	--	.007	---		.002				---
Occupation	.011	.012	.011	.025	.023	.026		.013				.026
Income	.016	.027	.017	.037	.023	.026		.021				.038
Perceived consumer effectiveness	.033*	.022	--	--	--	--		.018				--
Aggression	.016	.018	.028*	.002	.011	.005		.021				--
Depression	.007	.012	.015	--	.052	.039		.014				--
Self-esteem	.008	.027	.014	.017	.040	.005		.019				--
Sentience	.009	.022	.008	.079	.019	.035		.018				--
Tolerance	.012	.034	.017	.010	.065*	.006		.061*				--
Understanding	.016	.012	.015	.069	.021	.018		.043				--
Anxiety	.005	.024	.010	.072	.030	.020		.024				--
Desirability	.007	.010	.101	.064	.046	.033		.016				--
Dominance	.101	.013	.016	.029	.046	.033		.018				--
Harmavoidance	.008	.026	.007	.112*	.046	---		.029				--
Play	.007	.034	.015	.103	.023	.020		.045				--
Rebelliousness	.010	.014	.023	.054	.060	.063*		.015				--

*AID split on this variable.

TABLE 4
 THE PREDICTORS USED IN MCA RUNS AS THEY RELATE
 TO THREE BEHAVIORAL MEASURES
 OF ECOLOGICAL CONCERN

Dependent Variable		
Nonphosphate Laundry Product	Returnable Bottles	Existence of an Unusual Shopping Pattern
Perceived Consumer Effectiveness	Perceived Consumer Effectiveness	Perceived Consumer Effectiveness
Aggression	Aggression	Aggression
Self-esteem
Tolerance	Tolerance	Tolerance
Understanding	Understanding	...
Dominance	Dominance	...
Harmavoidance	...	Harmavoidance
Play	Play	Play
Income	Income	Income
Presence of children	Presence of children	...
...	Depression	...
...	Anxiety	...
...	...	Age of wife
...	...	Education of husband
...	...	Rebelliousness

TABLE 5

MCA RESULTS WITH THE USE OR NONUSE OF NONPHOSPHATE
LAUNDRY PRODUCTS AS DEPENDENT VARIABLES*

Factor	Number of Cases	Unadjusted Mean	Unadjusted Coefficient	Adjusted Mean	Adjusted Coefficient
<u>Presence of children</u>					
None under 19	133	.368	.001	.366	-.001
Under 6 only	75	.347	-.020	.343	.007
6-12 only	55	.345	-.021	.383	.016
13-18 only	71	.380	.014	.358	-.009
Under 6 and 6-12	62	.242	-.125	.223	-.144
6-12 and 13-18	76	.381	.015	.398	.011
All 3 age groups	20	.700	.333	.683	.316
Under 6 and 13-18	4	.500	.133	.631	.264
Missing data	3	.667	.300	.658	.291
Eta ² = .03					
<u>Income</u>					
Under \$5,000	50	.320	-.047	.351	-.016
\$5,000 to \$6,999	112	.330	-.036	.359	-.008
\$7,000 to \$9,999	181	.364	-.002	.367	.000
\$10,000 to \$14,999	110	.373	.006	.353	-.014
\$15,000 and over	31	.516	.149	.469	.102
Missing data	15	.467	.099	.365	-.002

Eta² = .01

Table 5 (Continued)

Factor	Number of Cases	Unadjusted Mean	Unadjusted Coefficient	Adjusted Mean	Adjusted Coefficient
<u>Perceived Consumer Effectiveness</u>					
1	19	.526	.167	.521	.154
2	60	.183	-.181	.190	-.176
3	35	.114	-.252	.155	-.212
4	202	.297	-.0	.294	-.073
5	179	.536	.176	.531	.164
Missing data	4	.500	.133	.506	.139
Eta ² = .09					
<u>Aggression</u>					
1	51	.431	.065	.381	.014
2	71	.423	.056	.360	-.007
3	82	.341	-.025	.308	-.059
4	89	.348	-.018	.323	-.044
5	55	.345	-.021	.335	-.032
Missing data	151	.351	-.016	.435	.068
Eta ² = .01					
<u>Self-esteem</u>					
1	71	.366	-.001	.356	-.011
2	73	.329	-.038	.383	.016
3	58	.414	.047	.436	.069
4	66	.439	.073	.414	.047
5	80	.338	-.029	.303	-.064
Missing data	151	.351	-.016	.351	-.016
Eta ² = .01					

Table 5 (Continued)

Factor	Number of Cases	Unadjusted Mean	Unadjusted Coefficient	Adjusted Mean	Adjusted Coefficient
<u>Tolerance</u>					
1	58	.362	-.005	.416	.049
2	66	.333	-.033	.350	-.017
3	90	.322	-.045	.318	-.049
4	74	.432	.066	.414	.047
5	60	.433	.067	.393	.026
Missing data	151	.35.	-.016	.351	-.016
Eta ² = .01					
<u>Understanding</u>					
1	70	.214	-.152	.299	-.078
2	75	.382	.020	.371	.003
3	78	.449	.082	.425	.058
4	71	.423	.056	.415	.048
5	54	.389	.022	.363	-.004
Missing data	151	.351	-.016	.351	-.016
Eta ² = .02					
<u>Dominance</u>					
1	68	.412	.045	.435	.069
2	89	.382	.015	.401	.034
3	61	.311	-.055	.331	-.035
4	68	.382	-.016	.324	-.042
5	62	.371	.004	.362	-.005
Missing data	151	.351	-.016	.351	-.016
Eta ² = .00					

Table 5 (Continued)

Factor	Number of Cases	Unadjusted Mean	Unadjusted Coefficient	Adjusted Mean	Adjusted Coefficient
<u>Harmavoidance</u>					
1	68	.353	-.014	.347	-.020
2	79	.418	.051	.424	.057
3	71	.380	.014	.399	.032
4	76	.447	.081	.446	.079
5	54	.222	-.144	.199	-.168
Missing data	151	.351	-.016	.351	-.016
Eta ² = .02					
<u>Play</u>					
1	66	.333	-.033	.321	-.046
2	86	.465	.098	.458	.091
3	88	.318	-.049	.346	-.021
4	59	.339	-.028	.313	-.054
5	49	.408	.041	.421	.054
Missing data	151	.351	-.016	.351	-.016
Eta ² = .01					

TABLE 6
MCA Results with the Use or Nonuse of Returnable
Bottles as Dependent Variable

Factor	Number of Cases	Unadjusted Mean	Unadjusted Coefficient	Adjusted Mean	Adjusted Coefficient
<u>Presence of children</u>					
None under 19	129	.496	-0.93	.485	-.104
Under 6 only	75	.520	-0.69	.512	-.077
6-12 only	55	.655	.065	.677	.088
13-18 only	71	.634	.045	.620	.031
Under 6 and 6-12	61	.639	.050	.612	.023
6-12 and 13-18	76	.658	.069	.689	.100
All 3 age groups	20	.750	.161	.764	.178
Under 6 and 13-18					
Missing data					
Eta ² = .03					
<u>Income</u>					
Under \$5,000	50	.580	-.009	.636	.047
\$5,000 to \$6,999	109	.624	.035	.625	.036
\$7,000 to \$9,999	179	.564	-.025	.563	-.026
\$10,000 to \$14,999	110	.600	.011	.593	.004
\$15,000 and over	31	.645	.056	.597	.008
Missing data	15	.467	-.122	.439	-.150
Eta ² = .01					

TABLE 6 (Continued)

Factor	Number of Cases	Unadjusted Mean	Unadjusted Coefficient	Adjusted Mean	Adjusted Coefficient
<u>Perceived Consumer Effectiveness</u>					
1	19	.579	-.010	.514	-.075
2	60	.567	-.022	.540	-.049
3	35	.429	-.161	.453	-.136
4	202	.559	-.030	.561	-.028
5	178	.663	.074	.672	.083
Missing Data	5	.568	-.021	.578	-.011
2					
Eta ² = .02					
<u>Aggression</u>					
1	50	.640	.051	.511	-.078
2	71	.592	.002	.474	-.115
3	81	.630	.041	.500	-.089
4	89	.629	.040	.470	-.119
5	54	.556	-.034	.431	-.158
Missing Data	149	.537	-.052	.847	.258
2					
Eta ² = .01					
<u>Depression</u>					
1	114	.640	.051	.621	.032
2	85	.588	-.001	.575	-.014
3	71	.620	.031	.650	.061
4	47	.596	.007	.612	.023
5	28	.571	-.018	.584	-.005
Missing Data	149	.537	-.052	.537	-.052
2					
Eta ² = .01					

TABLE 6 (Continued)

Factors	Number of Cases	Unadjusted Mean	Unadjusted Coefficient	Adjusted Mean	Adjusted Coefficient
<u>Tolerance</u>					
1	58	.569	-.020	.609	.020
2	64	.625	.036	.652	.063
3	90	.600	.011	.604	.015
4	73	.616	.027	.621	.032
5	60	.650	.061	.570	-.019
Missing Data	149	.537	-.052	.537	-.052
Eta ² = .01					
<u>Understanding</u>					
1	69	.507	-.082	.558	-.031
2	75	.667	.078	.661	.072
3	77	.558	-.031	.536	-.053
4	70	.671	.082	.676	.087
5	54	.667	.078	.637	.048
Missing Data	149	.537	-.052	.537	-.052
Eta ² = .02					
<u>Anxiety</u>					
1	56	.536	-.053	.532	-.057
2	74	.649	.060	.639	.050
3	69	.565	-.024	.551	-.038
4	74	.676	.087	.670	.081
5	72	.611	.022	.644	.055
Missing Data	149	.537	-.052	.537	-.052
Eta ² = .01					

TABLE 6 (Continued)

Factors	Number of Cases	Unadjusted Mean	Unadjusted Coefficient	Adjusted Mean	Adjusted Coefficient
<u>Dominance</u>					
1	68	.588	-.001	.593	.004
2	87	.621	.032	.633	.044
3	61	.508	-.081	.505	-.084
4	67	.731	.142	.710	.121
5	62	.597	.008	.600	.011
Missing Data	149	.537	-.052	.537	-.052
Eta ² = .02					
<u>Play</u>					
1	65	.631	.042	.608	.019
2	86	.523	-.066	.514	-.075
3	86	.674	.085	.668	.079
4	59	.593	.004	.627	.038
5	49	.653	.064	.668	.079
Missing Data	149	.537	-.052	.537	-.052
Eta ² = .01					

TABLE 7

MCA Results with the Existence or Nonexistence of an
Unusual Shopping Pattern as Dependent Variable

Factors	Number of Cases	Unadjusted Mean	Unadjusted Coefficient	Adjusted Mean	Adjusted Coefficient
<u>Age of wife</u>					
Under 25	15	.200	.050	.160	.010
25 - 34	152	.164	.014	.139	-.011
35 - 44	134	.164	.014	.141	-.009
45 - 54	124	.137	-.013	.129	.021
55 and over	61	.066	-.084	.149	-.001
Missing data	3	.667	.517	.174	.024
Eta ² = .02					
<u>Education of husband</u>					
Some/all elementary school	88	.058	-.091	.059	-.091
Some/all high school	278	.151	.001	.147	-.003
Some/all university or college	100	.230	.080	.235	.085
No husband	21	.048	-.102	.082	-.068
Missing data	3	.667	.517	.667	.517
Eta ² = .04					

TABLE 7 (Continued)

Factors	Number of Cases	Unadjusted Mean	Unadjusted Coefficient	Adjusted Mean	Adjusted Coefficient
<u>Income</u>					
Under \$5,000	50	.060	-.090	.132	-.018
\$5,000 to \$6,999	118	.148	-.001	.190	.040
\$7,000 to \$9,999	176	.135	-.013	.144	-.005
\$10,000 to \$14,999	110	.200	.050	.172	.022
\$15,000 and over	29	.138	-.012	.015	-.135
Missing data	15	.267	.177	.091	-.059
Eta ² = .01					
<u>Aggression</u>					
1	50	.082	-.068	.040	-.110
2	70	.071	-.078	.071	-.084
3	81	.185	-.036	.195	.045
4	88	.170	.020	.175	.025
5	55	.236	.087	.234	.084
Missing Data	145	.145	-.005	.144	.006
Eta = .02					
<u>Tolerance</u>					
1	57	.088	-.062	.112	-.038
2	64	.141	-.009	.122	-.028
3	90	.111	-.039	.130	-.030
4	72	.167	.017	.142	-.008
5	60	.267	.117	.282	.132
Missing data	145	.145	-.005	.145	-.005
Eta = .02					

TABLE 7 (Continued)

Factors	Number of Cases	Unadjusted Mean	Unadjusted Coefficient	Adjusted Mean	Adjusted Coefficient
<u>Harmavoidance</u>					
1	68	.191	.041	.159	.009
2	77	.195	.045	.175	.025
3	71	.127	-.023	.140	-.010
4	73	.137	-.013	.167	.016
5	54	.093	-.057	.107	-.043
Missing data	145	.145	-.005	.145	-.005
Eta ² = .01					
<u>Play</u>					
1	64	.156	.006	.163	.013
2	84	.083	-.067	.098	-.052
3	87	.207	.057	.206	.056
4	59	.136	-.014	.019	-.031
5	49	.184	.034	.175	.025
Missing data	145	.145	-.005	.145	-.005
Eta ² = .01					
<u>Rebelliousness</u>					
1	54	.167	.017	.183	.033
2	84	.095	-.054	.118	-.032
3	56	.107	-.042	.089	-.061
4	84	.179	.029	.190	.040
5	65	.216	.065	.176	.026
Missing data	145	.145	-.005	.149	-.005
Eta ² = .01					

TABLE 7 (Continued)

Factors	Number of Cases	Unadjusted Mean	Unadjusted Coefficient	Adjusted Mean	Adjusted Coefficient
<u>Perceived Consumer Effectiveness</u>					
1	19	.264	.114	.257	.107
2	58	.052	-.098	.034	-.116
3	34	.000	-.150	.062	-.088
4	199	.126	-.024	.131	-.019
5	178	.225	.075	.214	.064
2					
Eta = .04					

PERSONALITY AND CONSUMER BEHAVIOR: EXTENSIONS

Masao Nakanishi
University of California at Los Angeles

In a recent review of past and current research in personality and consumer behavior, Kassirjian noted (1971):

Purchasing behavior, media choice, innovation, segmentation, fear, social influence, product choice, opinion leadership, risk taking, attitude change, and almost anything else one can think of have been linked to personality.

The four papers presented today by Landon (1972), Bither and Dolich (1972), Gardner (1972), and Kinneer, Taylor and Ahmed (1972) also reflect the special fascination toward the personality concept held by the researchers in consumer behavior. Kassirjian claimed that the proportion of the total variance in the criterion variables explained by personality construct measures has been typically between 5 and 10 per cent in most studies. The papers presented in this session, with the exception of the Bither and Dolich paper, report similar lack of explanatory power by personality measures.

The papers by Landon and Gardner both deal with the individual differences in need for achievement and their relationship with product choice. As he acknowledges, Gardner's study must be treated as exploratory due to the smallness of the sample imposed by the measurement technique--Thematic Apperception Test. In the Landon paper, the operational problems of measurement are somewhat alleviated by the use of paper and pencil tests, but his conclusions are in doubt because of major flaws in statistical analysis. The Kinneer, Taylor and Ahmed paper tried to explain ecologically constructive behavior on the basis of socio-economic and personality variables, but found instead that an attitudinal variable--perceived consumer effectiveness--was the most significant variable. Only in the technically competent paper of Bither and Dolich, do we find some positive results on the relationship between personality and store choice behavior.

Both Kassirjian and Jacoby (1971) have attributed the lack of explanatory power to two common failings of past studies, i.e., 1) the slipshod adaptation of instruments designed for clinical uses, and 2) the lack of specific logical hypotheses to be tested. Though fully in agreement with those conclusions, skepticism remains if significant improvements in future research findings will result merely by more careful selection of test instruments and more experimental research designs to test specific hypotheses. The low explanatory power of individuals' personality characteristics, doubtlessly a pervasive force in shaping one's behavior, may have stemmed in part from naive conceptualization of the relationships between personality and consumer behavior often held by researchers in the field.

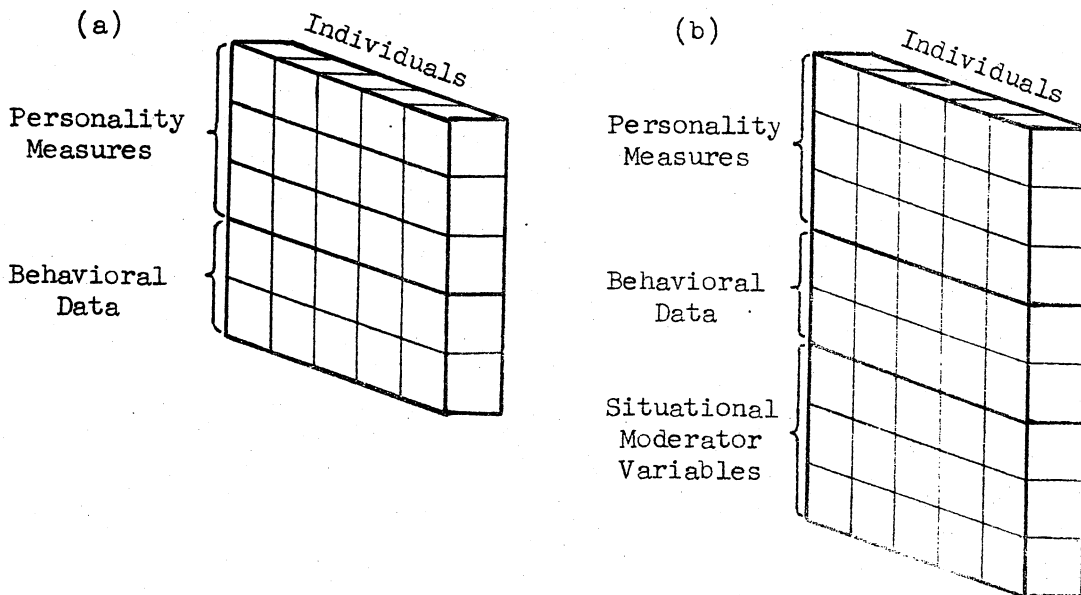
One of the naive concepts is that somehow an individual's personality has direct effects on his product and brand choices, much in the classical analysis of variance sense. The papers by Landon and Gardner appear to share this concept. This viewpoint is prompted by some researcher's desire to find easy methods of identifying potential customers and/or segmenting the market, but its naivete has become more and more apparent with recent research in consumer decision making. We now know that an individual's decision to purchase a

product or a brand is only a final result of a series of sub-decisions, such as decisions on store choice and choice criteria for each product category. Only recently a number of studies began exploring personality effects on the consumer decisions other than those of product/brand choice. Mark Alpert (1972), for example, investigated personality effects on choice criteria, while the Bither and Dolich paper in this session focused on the store selection decisions.

It may be argued that, if intermediate consumer decisions leading to the final product/brand choice are influenced by personality characteristics, then product/brand choice decisions must also be some function of personality. It would be difficult to refute this argument, but the functional relationships between personality and product/brand choice may not be so simple as to be amenable to analysis by standard statistical techniques which are often designed to measure linear relationships. The desire to find quick segmentation variables notwithstanding, it will be more fruitful to investigate the personality effects on each identifiable stage of the consumer decision making process separately. When we know more of personality effects on various stages, we may be able to synthesize the findings into a cohesive theory of personality effects on consumer behavior. Though studies by Alpert and Bither and Dolich cover only a small ground compared with the vast area of consumer behavior research, their effort should be considered as the beginning of a new and promising area of personality research.

Another naive view, often held unconsciously, is that of stationality of personality effects on individual's behavior. Most previous studies, including those reported in this session, may be called "one-shot" studies, the main objective of which is the search for behavioral correlates among personality construct measures over a sample of individuals taken at one point in time. Figure 1 (a) depicts the data set typically used. The behavioral data could be

Figure 1



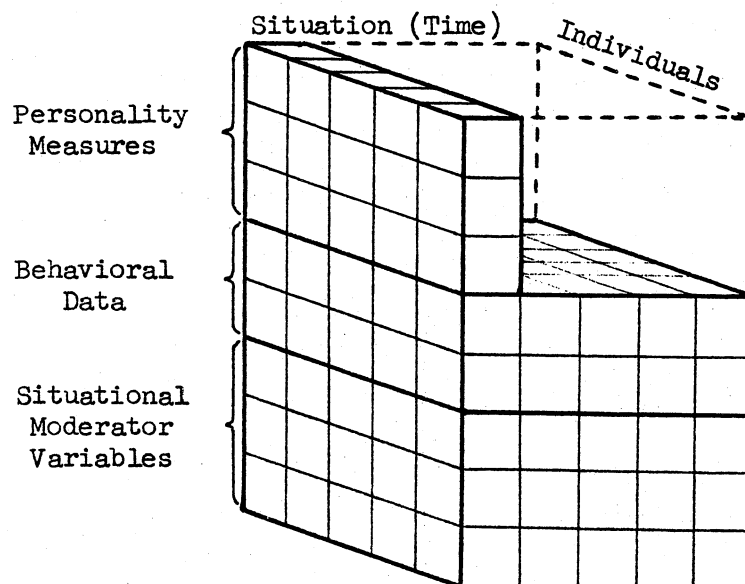
actual or hypothetical choices, intentions, or preferences for any alternatives, not necessarily products or brands. Those studies assume by design that an individual with a given personality configuration always behaves one way at different time points. This of course represents an extremely static view of human behavior and runs counter to what we know of its adaptive nature.

Kassarjian suggested a definition of personality as the "generalized patterns of response or modes of coping with the world." Those with the static viewpoint would interpret this as meaning the consistency of an individual's behavior over variety of situations, but personality may also be related to the manner the individual adjusts to the changes in his environment. In fact, it may be more correct to conceive personality as a moderator variable whose function is to moderate the effects of environmental changes on individual's behavior. This dynamic concept of personality effect have not been taken seriously by those who are interested in personality research.

Some researchers, for example, Brody and Cunningham (1968) and Fry (1971), tried to alleviate the lack of dynamism by incorporating situational and/or moderator variables in their research design and achieved a limited degree of success. Their objective was to isolate the main effects of personality on consumer behavior by explicitly adopting classical analysis of variance models. Figure 1 (b) gives the type of data set used by those studies. This research design implicitly assumes that moderator variables, such as perceived risk and specific self confidence, adequately summarizes feedbacks from past behavioral and situational events. However, as research in personality and persuasibility attests, personality may also affect over-time changes in moderator variables. It would, then, be more meaningful and potentially more powerful in predicting behavior if one can find out how moderator variables are formed and/or changed and how they interact with individual's personality. The one-shot research designs of Figure 1 tell us nothing of this type of interaction.

In order to investigate the full effects of personality on consumer behavior, it becomes necessary to obtain behavioral data and personality construct measures for each individual over different situations, accompanied by the data on situational and moderator variables. The data thus created would look as in Figure 2.

Figure 2

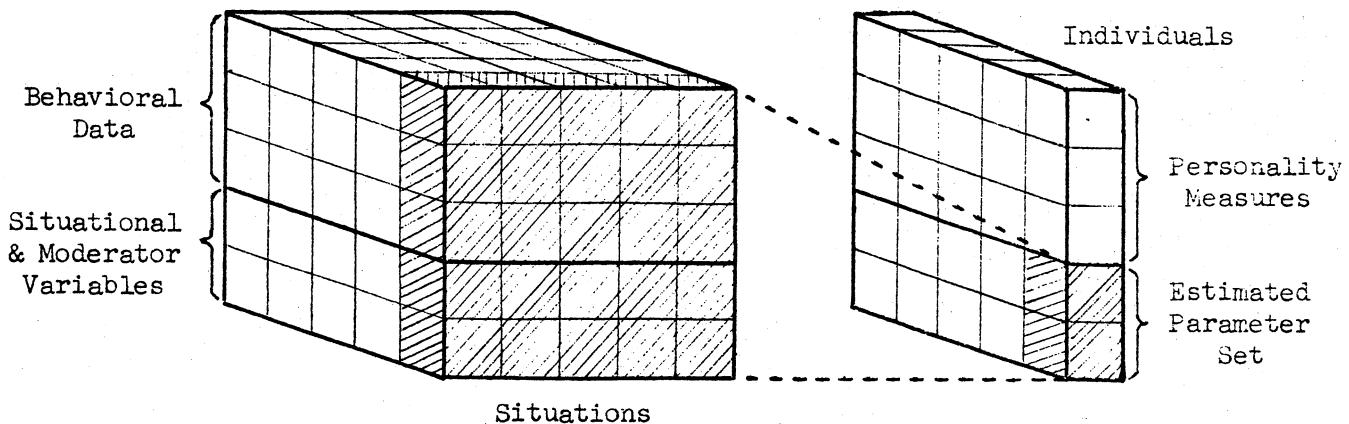


Analyzing data sets of this type would be a difficult undertaking, however. One might think of techniques analogous to a combination of cross-sectional and time series analysis in econometrics, but analysis would be made complex by interaction effects among predictor variables and lagged effects expected from the relationships between moderator variables and behavioral and situational variables. What is needed is an entirely new way of simultaneously analyzing those three types of variables.

An approach probably comes closest to the concept of personality as a moderator variable between behavior and environmental changes. Instead of combining personality construct measures and situational variables as predictors in a single model, one may construct a primary model which links behavioral data with situational and moderator variables at the individual level. This model is fitted to the data individual by individual and its parameters are estimated for each individual. Then secondary models may be constructed to specify the relationships between the estimated parameter set and personality construct measures. This two-step procedure is illustrated in Figure 3.

Since the two-step approach eventually links behavior as a function of personality, it may seem that it is no different from the conventional methods of analysis. But the difference is more than superficial. First, there is no need to restrict the primary and secondary models to naive linear forms. Fitting of non-linear models in the first stage followed by a multivariate analysis such as canonical correlation in the second stage is a perfectly permissible combination. It is conceivable to use as complex a model as the Howard-Sheth model in the first stage. Secondly, the two-step approach would become more attractive as the number of predictor variables (personality, situational, moderator) becomes large. The first stage becomes in fact a data reduction procedure, if the number of parameters to be estimated is small relative to the number of predictors.

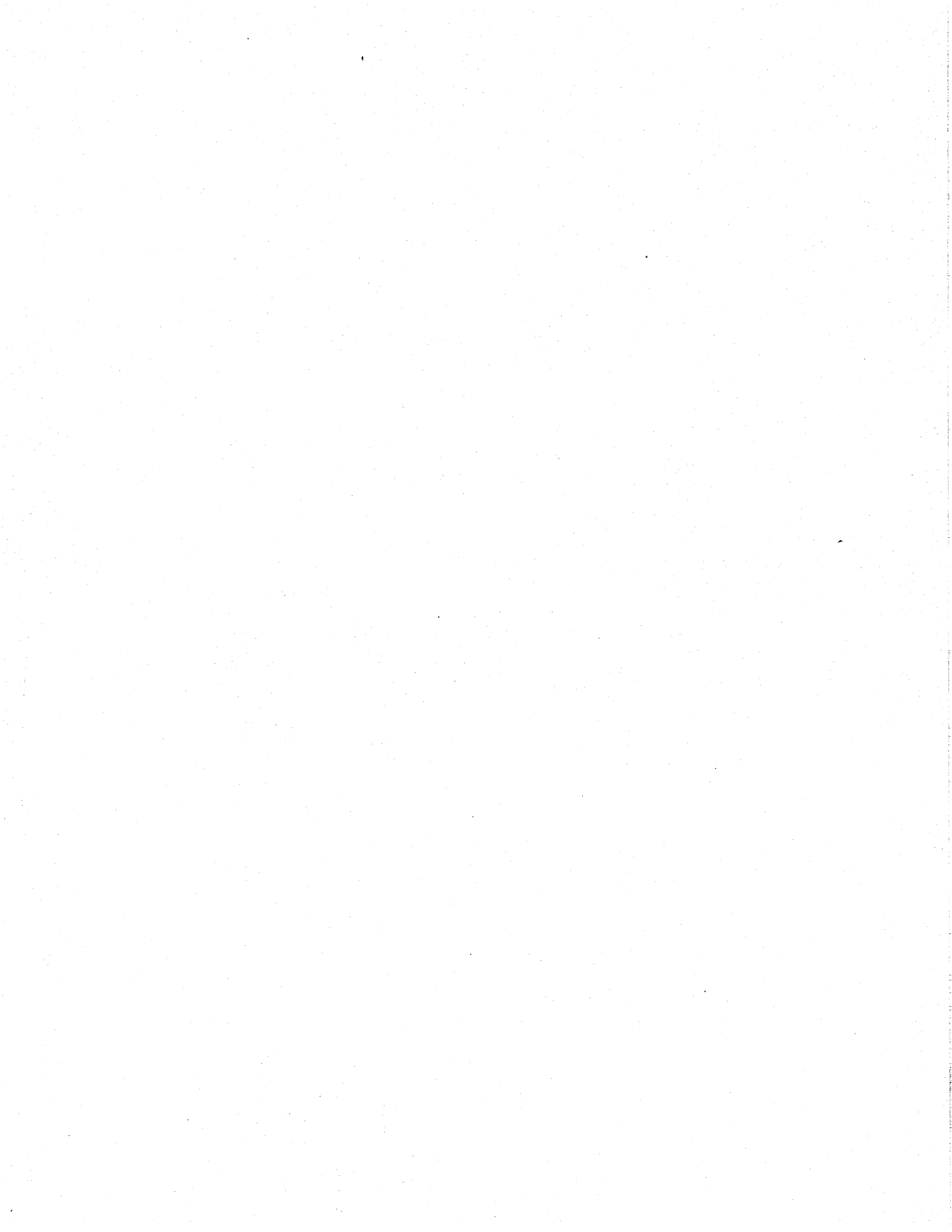
Figure 3



The purpose of this discussion was to show how naive concepts of personality effects on consumer behavior tend to obscure the areas of research which are most interesting and rewarding. To a degree, the naivete may arise from the limitations imposed by conventional techniques of statistical analysis. Although it is a truism that only clear conceptualization of relationships among variables under study helps the researcher determine the data to be obtained and the analytical techniques to be used, it is also true that we tend to gravitate toward an easy habit of conceptualizing everything in terms of simplistic linear relationships. The more pervasive the effect of personality, the more critical it becomes to clarify the myriad of relationships among many layers of variables interacting with each other since such complex interactions may not reveal themselves as simple linear relationships. In order for us to be able to develop a meaningful theory of personality effects it will be necessary to break out of the mold into which naive conceptualization has forced us.

REFERENCES

- Alpert, M. I. Personality and the Determinants of Product Choice. Journal of Marketing Research, 1972, 9, 89-92.
- Bither, S. W. & Dolich, I. J. Personality as a Determinant Factor in Store Choice. Paper presented at the Third Annual Conference of the Association for Consumer Research, Chicago, Illinois, November 1972.
- Brody, R. P. & Cunningham, S. M. Personality Variables and the Consumer Decision Process. Journal of Marketing Research, 1968, 5, 50-57.
- Fry, J. N. Personality Variables and Cigarette Brand Choice. Journal of Marketing Research, 1971, 8, 298-304.
- Gardner, D. M. An Exploratory Investigation of Achievement Motivation Effects on Consumer Behavior. Paper presented at the Third Annual Conference of the Association for Consumer Research, Chicago, Illinois, November 1972.
- Jacoby, J. Personality and Innovation Proneness. Journal of Marketing Research, 1971, 8, 244-247.
- Kassarjian, H. H. Personality and Consumer Behavior: A Review. Journal of Marketing Research, 1971, 8, 409-418.
- Kinnear, T., Taylor, J. R., & Ahmed, S. A. Socio-economic and Personality Correlates of Ecologically Constructive Purchasing Behavior. Paper presented at the Third Annual Conference of the Association for Consumer Research, Chicago, Illinois, November 1972.
- Landon, E. L. A Sex-Role Explanation of Purchase Intention Differences of Consumers Who Are High and Low in Need for Achievement. Paper presented at the Third Annual Conference of the Association for Consumer Research, Chicago, Illinois, November 1972.



CONSUMERISM, CONSUMER EXPECTATIONS, AND PERCEIVED PRODUCT PERFORMANCE

Rolph E. Anderson
Old Dominion University
and Joseph F. Hair, Jr.¹
The University of Mississippi

In America's era of agriculture which preceded the Industrial Revolution, the family was both a producing and consuming unit. Domestic arts furnished the major share of consumption goods, and the limited manufacturing was largely accomplished by local craftsmen. An illustration of the degree of self-sufficiency of the pioneer family is found in the 1787 newspaper publication of a farmer's letter, which read: "At this time my farm gave me and my whole family a good living on the product of it, and left me one year with 150 silver dollars, for I never spent more than ten dollars a year which was for salt, nails, and the like. Nothing to eat, drink, or wear was bought, as my farm provided all" (Early, 1898).

Consumer Expectations

Consumer expectations in the marketplace were not a serious problem when the family itself produced most of the essentials for everyday life. With home produced goods, the consumer of the goods had firsthand knowledge of their quality and workmanship. As for the few items that were purchased, the buyer had a much better understanding of the techniques of production and the tests of quality than do most consumers today. Even though the problem of recognizing the comparative quality of goods was minimal when goods were produced by those who would use them or by local tradesmen who were personally known by their customers, this does not imply that goods were necessarily of better quality than today. For example, homespun woolen cloth was typically coarse, rough, and uneven. It could not begin to match the fine texture made possible by modern manufacturing techniques. But, whatever the quality, it was known by the purchaser beforehand, so he knew exactly what to expect from the product. Thus, he was seldom disappointed.

Rising Expectations

American business is generally considered to be doing a better job for consumers than ever before. In the words of Otto Kleppner, "...Today's average refrigerator has a far better refrigerant, a better motor, better insulation, and larger storage space than the costliest model of twenty years ago. Canned soups today have better enriched recipes than when you were a youngster, and are offered in greater variety. Today's tires give over four times the mileage per dollar (aside from excise taxes) of those of twenty years ago, in addition to giving you a smoother and safer ride" (Kleppner, 1970). In further support of this contention, Time magazine (December 12, 1969, p. 92) reported that the average buyer probably gets more value per dollar spent in a current mail order house catalogue than in an edition fifty years ago. Why, then, do we have so many complaints against products and the rising demand for consumer protection legislation?

Arjay R. Miller, former Ford vice-chairman and now dean of Stanford's Graduate School of Business, attributes the growing consumer irritation with the slightest flaw to the "phenomenon of rising expectations" (Business Week,

1969). Largely due to the increase in consumer affluence and sophistication, the threshold of acceptable performance is rising. Younger, wealthier, better educated, and more sophisticated consumers seem to be less tolerant of gaps between promotional promise and product performance.

Another contributor to rising consumer expectations in general has been the remarkable space flights. Many Americans reason that if we can put a man on the moon, why can we not find a way to eliminate pollution and poverty and make products that do not fail? The mass communications achieved by television and the transistor radio and the great mobility of people have carried the message of hope and high expectations worldwide. No one can predict human expectations. They are infinitely elastic -- bounded only by individual imagination.

Sources of Information

The expectations consumers have regarding a particular product depend upon information gathered from a variety of sources. Past experience, promotional communications of sellers, and personal acquaintances are the most common sources of product information. If a product is purchased frequently, the consumer may have a satisfactory information source -- his previous experience with it. In such a case, he is able to judge, prior to purchase, the product's effectiveness in meeting his expectations, both functional and psychological. Conversely, when purchasing an untried product or brand of significant importance, the consumer may find a meaningful choice difficult because of the lack of information. Since he has not previously purchased this product, he must rely on information sources other than his personal experience. Another source may be his associates, but the limited accuracy and nontransferability of this experience to his own situation often sharply limits its value. Ratings by private, independent organizations are often regarded as excellent sources but tend to be used by those who least need help in purchasing decisions. For example, the most recent survey of subscribers to Consumers Union showed the median income is \$14,000 and 58 percent are college graduates (Herrmann, 1970). Thus, it appears that many consumers depend largely on one basic information source -- the company's promotion mix -- in forming their expectations regarding new or untried products.

Information Gap

Proponents of consumerism claim that corporate marketing programs are not providing the information necessary for meaningful choice, and point to inaccurate, misleading, and inadequate information as major reasons for the demand for more consumer protection legislation. As E. B. Weiss says: "When every detergent gets clothes whiter, brighter, cleaner, sweeter-smelling than any other; when every toothpaste is better than every other in preventing tooth decay; when every gasoline makes your car run better than any other, what is the consumer to believe since the claims can't all be true?" (Weiss, 1967).

Many consumer protectionists see improvements in promotion as the best means of spanning the expectations -- performance gap. Promotion, they insist, instead of adding confusion to an already complex marketplace, could be developing well-informed consumers by providing relevant and reliable information for use in making accurate buying judgments. This proposition maintains that well-informed individuals would not unknowingly buy inferior products, leading to dissatisfaction and demands for more consumer protection legislation. Given

enough relevant information, it is claimed, the consumer would be able to form realistic expectations about products and be able to protect himself from those evils from which the government is now trying to shield him.

Consumer Dissatisfaction

The Random House Dictionary states: "Dissatisfaction results from contemplating what falls short of one's wishes or expectations...." In much the same vein, Stanton has provided one of the best definitions of consumerism, as follows: (1) the reaction of consumers to their dissatisfactions and un-realized expectations and (2) their efforts to have these perceived injustices remedied (Stanton, 1971). Based on these definitions, one might hypothesize that consumer dissatisfaction and consumerism result from market offerings which fall short of consumer expectations. It may be that corporate promotional mixes are helping create excessively high expectations for products which result in consumer dissatisfaction upon purchase and use. This apparent gap between promise and performance may be largely responsible for rising support for legal enforcement of minimum performance standards. Buskirk and Rothe unequivocally declare: "It is this sense of frustration and bitterness on the part of consumers who have been promised much and have realized less, that may properly be called the driving force behind consumerism" (Buskirk and Rothe, 1970).

If business is providing communications that raise consumer expectations beyond a realistic level, then it is important to learn what kind of communication should be provided -- presumably still persuasive in nature but different in content. Knowledge about the effect of consumer expectations and perceived product performance on consumer satisfaction is vital because management can, within limits, influence consumer expectations and product performance.

Interrelated Variables

In order to meaningfully discuss interrelationships among expectations, perceived product performance, and consumer satisfaction, the three variables need clarification.

Expectations

Expectations have been described as "subjective notions of things to come" (Katona, 1958). An expectancy is a type of hypothesis formulated by the consumer, and his perception of product performance after purchase and use will serve to either confirm or reject this hypothesis (Engel et al, 1968). Expectations are confirmed when a consumer receives what he expects. Negative disconfirmation takes place when the product's actual or objective performance cannot match expectations for it. Positive disconfirmation may occur when objective product performance actually exceeds expectations.

Expectations may be created and strengthened by corporate promotional mixes, past experiences, opinions of friends and associates, impartial product rating services, or general aspiration levels. Writers frequently refer to consumer "attitudes" or "aspirations," or "images" of products, brands, or stores, and these terms may be considered implicit forms of expectations. For example, when a consumer expresses a favorable attitude toward a product, he may say: "That's a good product," but in translation, he means that he confidently expects that product to provide certain specific benefits. As Pro-

fessor Theodore Levitt has pointed out, people buy the "expectation of benefits" (Business Week, 1972). Consumers are not anxious to purchase new, sharper razor blades but faster, smoother shaves.

Product Perception

Perception can be thought of as the individual's mental impression of a stimulus object, in this case, a product. No two people perceive a product exactly alike because no two people have the same view of their environment. Perception has four major facets. First, it is selective. Since an individual cannot possibly be cognizant of all the stimulus objects within his perceptual field, he perceives selectively. Secondly, perception is organized in that it has meaning for the individual. Third, perception depends upon stimulus factors, e.g., advertising in consumer product evaluations. Fourth, perception is influenced by personal factors, i.e., what the individual brings to the situation. The consumer's view of the world, or his cognitive set, is formed over time, and reflects his physiological and psychological characteristics as well as his needs, moods, values, past experiences, and the nature of his environment (Robertson, 1970). In the present study, consumer perception of the product will be called "perceived product performance." "Actual" or "objective" product performance will be used in referring to reality or true product performance.

Consumer Satisfaction

No satisfactory literal definition has yet been developed for consumer satisfaction. However, inferring from our description of the source of dissatisfaction, one might postulate that satisfaction results when consumer expectations are matched (or exceeded) by perceived product performance. Consumer dissatisfaction, then, might be measured by the degree of disparity between expectations and perceived product performance.

Theories of Expectations

Except for the field of psychology, the literature of the social sciences contains few references to any systematic studies relating to the effects on people of disparity between their expectations and actual product performances. However, since consumer satisfaction is such a fundamental problem to human behavior, it is only logical that one would turn to psychology to develop a theoretical framework for the present investigation.

Assimilation (Cognitive Dissonance)

According to Festinger's theory of cognitive dissonance, an unconfirmed expectancy creates a state of psychological discomfort because the outcome contradicts the consumer's original hypothesis (Festinger, 1957). The theory suggests that any disparity between expectations and actual product performance will be minimized by the customer adjusting his perception of the product to become more consistent with his expectations. Consumers are continually receiving various types of information about products from their own experiences, associates, advertisements, and salesmen. These pieces of information are cognitions which consumers like to have consistent with one another (Holloway, 1967). When inconsistent information is received, the individual suffers mental discomfort which he attempts to reduce--often by changing or distorting one or more of the cognitions to make them more consonant with each other. The stronger the cognitive dissonance, the more motivated he is to alleviate tension by altering the cognitive elements (Brehm, 1962).

As applied to marketing, if there is disparity between expectations for a product and the actual performance of that product, the consumer is stimulated to lessen the psychological tension generated by changing his perception of the product so that it comes more into agreement with his expectations. Therefore, if this theory is true, the promotional mix for a product should substantially lead expectations above product performance to obtain higher consumer evaluations of the company's product. This concept is illustrated in Figure 1 by the dotted line which shows that perceived product performance is always between actual performance and consumer expectations, except when all three coincide. Considerable controversy and some disaffection with the theory of cognitive dissonance have developed in recent years due to the accumulation of an increasing amount of contradictory evidence (Chapanis and Chapanis, 1964; Feldman, 1966; Insko, 1967; and Rosenberg, 1965). A major criticism is that the theory assumes the individual does not learn from his purchasing mistakes. Instead, he actually increases the probability of repeating past mistakes through his efforts to reduce post-purchase dissonance by justification and rationalization of his decisions (Cohen and Goldberg, 1970).

Contrast

Even in the studies supporting assimilation theory, some individuals tend to shift their evaluations away from expectations stimulated by communications if inconsistent with reality (Cardozo, 1964; and Hovland, Harvey & Sherif, 1957). Contrast theory assumes that the consumer will magnify any disparity between the product received and the product expected; i.e., if the performance of the product falls short of his expectations, the customer will evaluate the product less favorably than if he had no prior expectations for it. When expectations are not matched by actual product performance, contrast theory presumes that the surprise effect or contrast between expectations and outcome will cause the consumer to exaggerate the disparity.

Contrast theory would predict consumer product perceptions as shown by the dashed line in Figure 1. It implies that slight understatement of the product's qualities in promotion might lead to higher customer satisfaction with the product. Obviously, the advertisements or other components of the promotional mix would not so understate the product's qualities that consumers by-pass it for another brand. Several studies lend support to the possible success of this promotional strategy (Diab, 1965; Freedman, 1964; Hovland, et al., 1957; Sherif & Hovland, 1961; Spector, 1956; and Whittaker, 1965).

Assimilation-Contrast

As its name implies, the assimilation-contrast approach combines the theories of assimilation and contrast. Hovland, Harvey, and Sherif have provided support for the contention that product performance differing only slightly from one's expectations tends to bring about adjustment of product perceptions toward expectations (assimilation effect), while large variances between one's expectations and actual product performance tend to be magnified or exaggerated (contrast effect) (Hovland, et al., 1957). The theory assumes that individuals have ranges or latitudes of acceptance, rejection, and neutrality. If the disparity between expectations and performance falls into the latitude of acceptance, there is strong probability of an assimilation effect. On the other hand, if the disparity falls into the consumer's latitude of rejection, a contrast effect is likely.

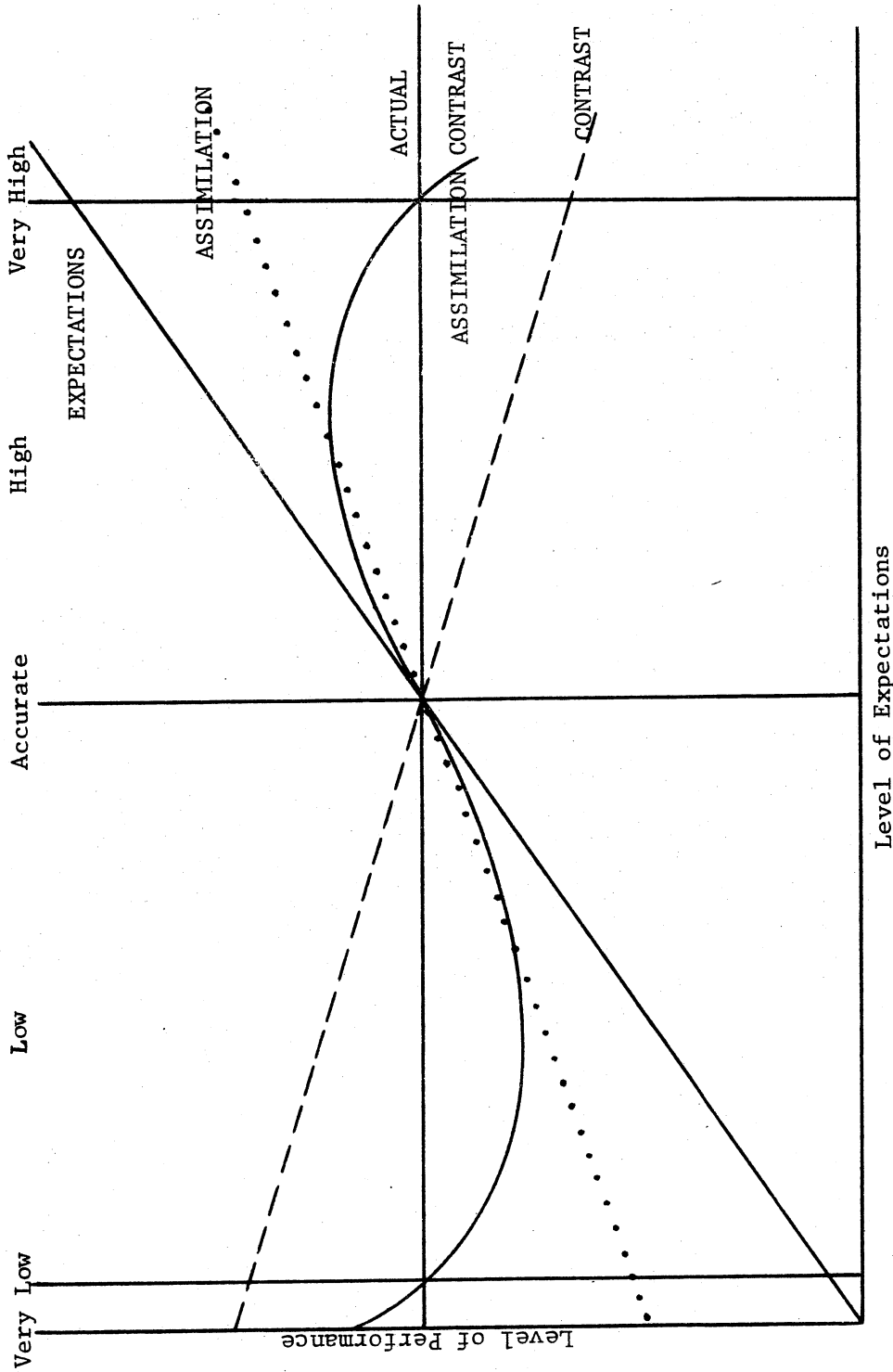


FIGURE 1. Theories of Disconfirmation of Expectations

Assimilation-contrast theory suggests that promotional messages ought to create expectations for the product as high as possible without creating a level of disparity between expectations and actual performance which exceeds the consumer's latitude of acceptance. In accord with assimilation-contrast theory, consumer perception of product performance would take the shape of the S-shaped curve in Figure 1. At low levels of disparity between expectations and product performance, consumer product perceptions tend to assimilate differences and draw near to expectations. However, as the disparity becomes large, it may reach the point where contrast effect takes over, and differences are magnified.

Hypotheses

Which, if any, of these theories best describes the true relationships between these important consumer variables? Four hypotheses may be considered, as follows:

1. Null -- product perceptions are not significantly different for various levels of expectations.
2. Assimilation -- product perceptions will vary directly with the level of expectations.
3. Contrast -- product perceptions will vary inversely with the level of expectations.
4. Assimilation-Contrast -- product perceptions will vary directly with expectations over a range around actual performance, but above and below this threshold, product perceptions will vary inversely with the level of expectations.

Methodology

Subjects (Ss) for the empirical investigation consisted of 144 volunteers from undergraduate marketing classes. No reward or incentive was offered for participation. However, after the experiment began, in order to increase involvement or commitment, subjects were told they could keep the product they were going to evaluate. The product selected for the experiment was a ballpoint pen, for several reasons: (a) students constitute a major market for this product; (b) students have been shown to be more interested in writing instruments than other miscellaneous products priced under two dollars (Cardozo, 1964); and (c) less variability of preference has been found among writing instruments than in most other product categories (Cardozo, 1964). All the pens were identical, unmarked ones selling at retail for about one dollar each.

Research Design

Using a 2 X 6 factorial design, the independent variable (expectations) was manipulated by randomly assigning subjects to one of six different conditions or levels of product information. As confirmed in pretesting, condition one (C₁) substantially understated the product features, C₂ slightly understated the features, C₃ depicted the product accurately, C₄ slightly overstated the quality of the product's features, and C₅ substantially overstated the product's features. C₀ provided no information about the product, but instead Ss received a communication unrelated to the experiment.

Half the Ss in each experimental condition were asked to complete a questionnaire regarding their expectations after reading the product information, but prior to seeing the product. This questionnaire acted as a "take-measure" to ensure that the different levels of product information were creating expectations in the right direction, with the desired degree of intensity. Remaining Ss were given a task unrelated to the experiment so that the time occupied by the two treatments was about equal. If the take-measure had proved reactive, responses of those Ss would have been discarded and only Ss not receiving the take-measure would have been included in the data analysis.

Each S was permitted to inspect and test the product for the same period of time, then record his reactions on a modified logarithmic product rating scale using dollars and cents distributed in small ranges from \$0.04 to \$64.00. These ratings were subsequently converted to integers by sequentially numbering the rows of the product rating scale. Three dependent measures were obtained. One dependent variable consisted of the mean of individual subject ratings on 15 visual features and performance characteristics of the ballpoint pen. The second variable was an overall rating by Ss on the pen's combined features. This was a weighted mean since each S could assign certain product features more importance than other features in determining the overall evaluation. Finally, Ss estimated the pen's price.

Results

As shown in Table 1, mean scores for all three dependent variables are assimilated toward expectations until reaching the "very high" level of expectations, which caused a sharp downturn in product ratings for all three measures.

One-way analysis of variance (treatment variable collapsed since Ss were not sensitized by the take-measure) revealed a main effect of conditions significant at the .01 level for all three dependent variables (product features, $F = 10.06$; combined characteristics, $F = 9.31$; and price, $F = 11.08$). Thus, Ss responded differently in their evaluations or perceptions of the product depending upon their level of expectations. There were no significant treatment main effects nor treatment by condition interaction effects for any of the dependent variables.

Relating Results to Theoretical Models

Ratings for each of the dependent variables were plotted by each of the six conditions, as illustrated for product features in Figure 2. Product ratings plotted on the vertical axis are the mean responses for all 24 Ss in each expectational condition. Expectations plotted on the horizontal axis are the mean expectations for the 12 Ss in each condition who were administered the take-measure. Inspection of the plotted data indicated conformity with assimilation theory until reaching C_5 , the "very high" level of expectations, which marked a decline in product evaluations for all three dependent variables in accord with assimilation-contrast theory. Not only were mean ratings in C_5 lower than in C_4 but, for both product features and combined characteristics, evaluations were lower in C_5 than in C_3 where accurate information was provided. C_0 was plotted separately (since expectations for these Ss were not manipulated) and proved significantly lower than C_3 .

TABLE 1. Mean Response by Condition and Treatment

Manipulation	C ₁ Very Low	C ₂ Low	C ₀ None	C ₃ Accurate	C ₄ High	C ₅ Very High
Expectations:						
Dependent Variable ^a						
A.	2.97	4.31	7.24	8.54	11.70	15.92
B.	3.25	4.25	8.42	8.92	12.42	16.67
C.	3.67	5.42	7.50	9.58	12.08	17.08
Product Ratings:						
Treatment 1 ^b						
A.	6.43	7.22	7.39	8.56	9.21	9.00
B.	6.58	7.83	8.00	9.00	9.75	9.25
C.	7.00	7.92	8.17	9.58	10.00	9.83
Treatment 2 ^b						
A.	6.22	6.84	6.96	8.94	10.05	8.38
B.	6.58	7.17	7.33	9.25	10.50	8.67
C.	6.92	7.92	8.50	9.58	10.75	10.08
Combined ^c						
A.	6.33	7.03	7.17	8.75	9.63	8.69
B.	6.58	7.50	7.67	9.13	10.13	8.96
C.	6.96	7.92	8.33	9.58	10.38	9.96

^aDependent variable A = Product Features, B = Combined Characteristics, C = Price

^bTreatment 1 = Take-measure, Treatment 2 = No take-measure

^cTreatment variable collapsed

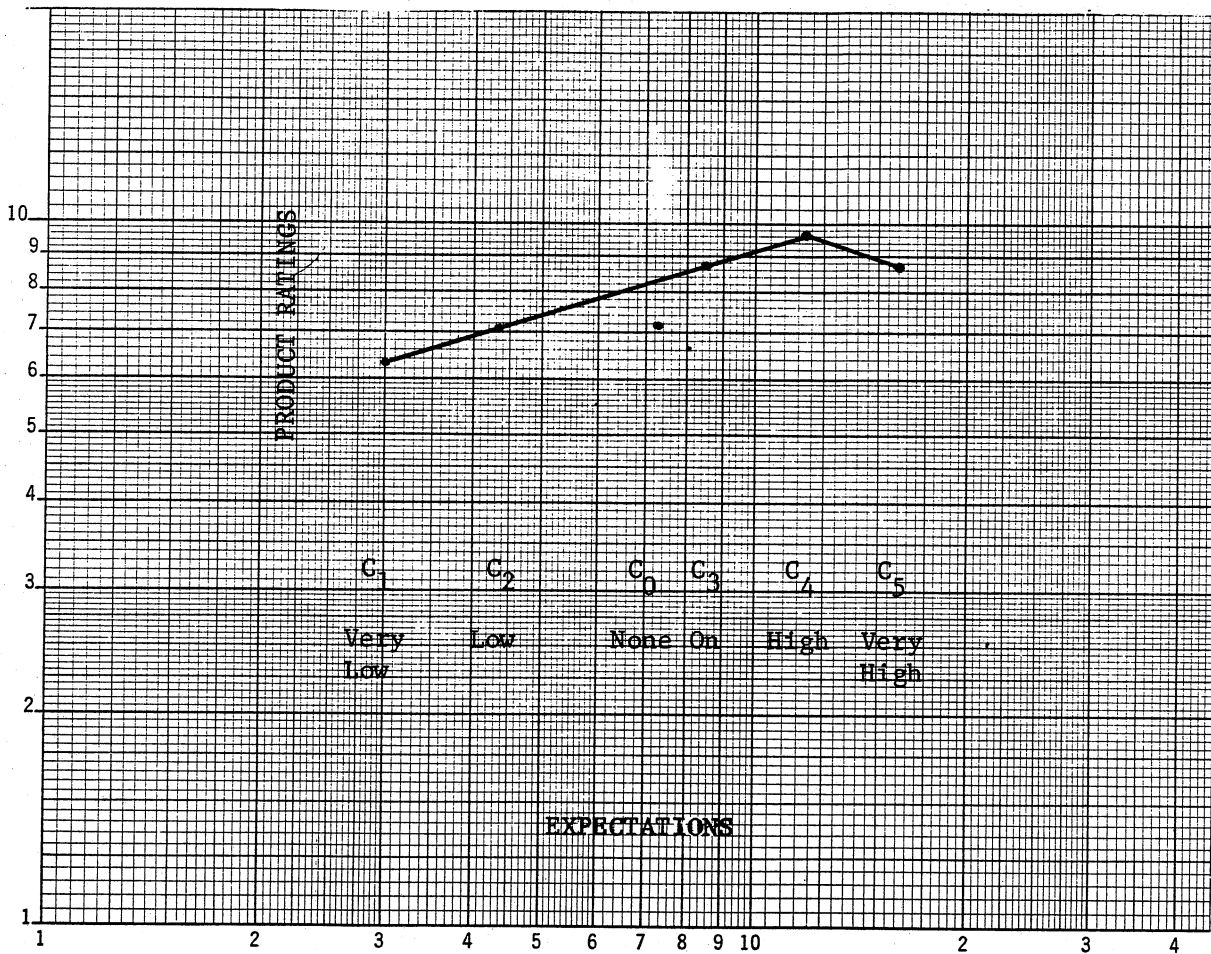


FIGURE 2. Plot of Mean Response by Condition:
Product Features

Linearity Tests

The tests for linearity, provided in Table 2, showed deviations from linearity to be highly significant for the relationship between expectations and product perceptions for each of the dependent variables. Significant deviations from linearity rejected both the null hypothesis of no effect of expectations and the assimilation hypothesis. Contrast theory which demands a negatively sloped relationship between expectations and product perceptions was quickly discarded because of the positive slope of the plotted data. Thus, the data best fit assimilation-contrast theory since product perceptions are assimilated toward expectations until the "very high" C₅ when contrast effect begins.

Although contrast effects did not appear in C₁, "very low" expectations, this result may be partially explained by "floor effects" which may have prevented manipulation of expectations far enough below the relatively low cost

TABLE 2

Deviation From Linearity: Relationship
Between Expectations and Product Perceptions

Sources	df	SS	MS	F
<u>Product Features</u>				
Between Groups	4	177.19		
Linear Regression	1	117.65		
Deviation From Linear Regression (D)	3	59.54	19.85	
Error Term (E)	115		3.88	
D/E				5.12**
<u>Combined Characteristics</u>				
Between Groups	4	189.75		
Linear Regression	1	117.17		
Deviation From Linear Regression (D)	3	72.58	24.19	
Error Term (E)	115		4.56	
D/E				5.30**
<u>Price</u>				
Between Groups	4	203.58		
Linear Regression	1	154.70		
Deviation From Linear Regression (D)	3	48.88	16.29	
Error Term (E)	115		3.91	
D/E				4.17**

**p < .01

product to elicit sufficient surprise or delight upon seeing and trying the pen. More complex products, where there is ambiguity and uncertainty in making judgments, may deliver different results as subjects may tend to rely more on the information provided. Olshavsky and Miller found assimilation theory supported in student evaluations of a reel-type tape recorder; i.e., overstatement for this complex, multidimensional product led to more favorable evaluations and understatement to less favorable evaluations (Olshavsky and Miller, 1972). To date, the only other published experiment in the marketing literature dealing with the effects of disconfirmed expectancies on consumer evaluations of products provided partial support for contrast theory (Cardozo, 1965). More research needs to be undertaken with a variety of products and services, e.g., ones requiring deep personal and financial commitments. Such items may yield substantially different results than those for less personal, lower cost, and less ego-related products and services. As an adjunct to such additional studies, it might be profitable to determine if there are significant differences between consumer reactions based on psychographic variables.

Conclusions

This study has demonstrated that consumer expectations can influence consumer perceptions of products. High expectations for product features and performance seem to generate, at least initially, higher ratings for the product up to a certain point, depending upon the product and its importance to the consumer. Beyond this critical threshold, however, exaggerated high expectations may be detrimental to product evaluations.

In sum, marketers ought to exercise caution in positioning the level of advertising and other promotional claims. Consumer expectations can profitably be led only so far for certain relatively simple and easily understood products. A tendency toward promotional hyperbole can be contagious among competing companies but can result in lower overall product evaluations and perhaps consumer dissatisfaction which may add other voices to the chorus calling for greater legislative control over the marketplace, especially on advertising and the promotional mix.

Footnotes

1. Rolph E. Anderson has several years business experience with consumer packaged goods manufacturers. His last position was New Product Development Manager for The Quaker Oats Company. He holds the B.A. and M.B.A. degrees from Michigan State University, and the Ph.D. degree from the University of Florida. He taught at the latter institution for two years.

Joseph F. Hair, Jr. has been employed in the Marketing Research Department of N. V. Philips' Gloeilampenfabrieken in Eindhoven, Netherlands and as a Systems Analyst for Food Fair Stores, Inc., Jacksonville, Florida. He holds the B.A., M.A., and Ph.D. degrees from the University of Florida, where he also taught for two years.

References

- Business Week. New Products: The Push Is On Marketing. (March 4, 1972), 72-77.
- Business Week. Why Detroit is Wary of the 1970's. (September 20, 1969), 110-19.
- Buskirk, R. H. & Rothe, J. T. Consumerism--An Interpretation. Journal of Marketing, 1970, 34, 61-65.

- Brehm, J. W. & Cohen, A. R. Explorations in Cognitive Dissonance. New York: John Wiley and Sons, 1962.
- Cardozo, R. N. An Experimental Study of Customer Effort, Expectations, and Satisfaction. Journal of Marketing Research, 1965, 2, 244-249.
- Cardozo, R. N. An Experimental Study of Customer Effort, Expectation, and Satisfaction. Unpublished doctoral dissertation, University of Minnesota, 1964.
- Chapanis, N. P. & Chapanis, A. Cognitive Dissonance: Five Years Later. Psychological Bulletin, 1964, 61, 1-22.
- Cohen, J. B. and Goldberg, M. E. The Dissonance Model in Post-Decision Product Evaluation. Journal of Marketing Research, 1970, 7, 315-21.
- Diab, L. N. Some Limitations of Existing Scales in the Measurement of Social Attitudes. Psychological Reports, 1965, 17, 427-30.
- Early, A. M. Home Life in Colonial Days. New York: The Macmillan Company, 1898.
- Engel, J. F., Kollat, D. T. & Blackwell, R. D. Consumer Behavior. New York: Holt, Rinehart and Winston, Inc., 1968.
- Feldman, S., ed. Cognitive Consistency: Motivational Antecedents and Behavioral Consequences. New York: Academic Press, 1966.
- Festinger, L. A Theory of Cognitive Dissonance. New York: Harper and Row, 1957.
- Freedman, J. L. Involvement, Discrepancy and Change. Journal of Abnormal and Social Psychology, 1964, 69, 290-95.
- Herrmann, R. O. Consumerism: Its Goals, Organizations and Future. Journal of Marketing, 1970, 34, 55-60.
- Holloway, R. J. An Experiment on Consumer Dissonance. Journal of Marketing, 1967, 31, 39-43.
- Hovland, C. I., Harvey, O. J. & Sherif, M. Assimilation and Contrast Effects in Reactions to Communication and Attitude Change. Journal of Abnormal and Social Psychology, 1957, 55, 244-52.
- Insko, C. A. Theories of Attitude Change. New York: Appleton-Century-Crofts, 1967.
- Katona, G. Business Expectations in the Framework of Psychological Economics (Toward a Theory of Expectations). In Mary Jean Bowman. Expectations, Uncertainty, and Business Behavior. New York: Social Science Research Council, 1958.
- Kleppner, O. The Role of Competitive Advertising. In Otto Kleppner (ed.). Exploring Advertising. Englewood Cliffs: Prentice-Hall, 1970, 9, 38-43.
- Olshavsky, R. W. & Miller, J. A. Consumer Expectations, Product Performance, and Perceived Product Quality. Journal of Marketing Research, 1972, 9, 19-21.
- Robertson, T. S. Consumer Behavior. Glenview: Scott, Foresman and Company, 1970.
- Rosenberg, M. J. When Dissonance Fails: On Eliminating Evaluation Apprehension from Attitude Measurement. Journal of Personality and Social Psychology, 1965, I, 28-42.
- Sherif, M. & Hovland, C. I. Social Judgment: Assimilation and Contrast Effects in Communication and Attitude Change. New Haven: Yale University Press, 1961.
- Spector, A. J. Expectations, Fulfillment, and Morale. Journal of Abnormal and Social Psychology, 1956, 52, 51-56.
- Stanton, W. J. Fundamentals of Marketing. New York: McGraw-Hill Book Company, 1971.
- Weiss, E. B. Advertising's Crisis of Confidence. Advertising Age, 1967, 38, 138-44.
- Whittaker, J. O. Attitude Change and Communication-Attitude Discrepancy. Journal of Social Psychology, 1965, 65, 141-47.

CONSUMER RESPONSE TO PRODUCT-LINE EXTENSIONS:
A CONSIDERATION FOR NEW PRODUCT PLANNING

J. Taylor Sims¹
University of South Carolina

If a household is familiar with a brand, it may be induced by a new flavor introduction to buy an extra unit sooner than it would otherwise. These effects would be part of the short-run impact of a new flavor introduction. Such an interim purchase would, however, be unlikely to affect their long-run purchases of the product category, with reference to the average number of items and volume attributed, in any substantial way. The average number of items within a grocery product class that contribute to the total assortment mix of products within a household should not change, given relatively stable socioeconomic conditions. Thus, it is proposed that the long-run effect of a new flavor introduction on a "repeat" buyer of the brand class to which the flavor belongs is inconsequential with reference to additional total brand volume that could be forthcoming.

If the above hypothesis can be substantiated then the impact of a product-line extension is primarily dependent on the retention of new buyers who currently either fulfill most of their requirements with competing items from other brands or have infiltrated from outside the product market. The ability of the product-line extension to lure and maintain new buyers, i.e., convert them into repeat purchasers, is the primary determinant of success.²

Methodology of the Study

The main emphasis of the study is analytical. It seems necessary to point this out because most of the previous studies dealing with brand loyalty have dealt at length on the strengths and weaknesses of methodological techniques instead of the resulting analyses (Brown, 1962; Cunningham, 1961; Fourt & Woodlock, 1966; Frank, 1962; Howard, 1965; Kuehn, 1962; Massy, 1967; and Parfitt & Collins, 1968). In addition, these investigations have examined product loyalty primarily at the brand level with little or no attempt to look at type compositions within brand structure. In contrast, the emphasis of this study is on the analysis of type as well as brand loyalty.

In order to test the hypothesis of this study, data from the Market Research Corporation of America's National Consumer Panel on the introduction of new types within a particular brand of a specific product class were examined. Other product classes were made available, on a limited basis, for comparative analyses.

The National Consumer Panel consists of 7,500 households stratified to represent the demographic structure of the total United State population. Continuous purchase data are collected by means of a diary which is filled out by the consumer and returned weekly for processing into cumulative computerized data files. These data can be classified in numerous ways (Market Research Corporation of America, 1970). For the purpose of this study data were classified in the following manner:

1. The number of total households purchasing a product class on a calendar quarter to quarter basis;
2. The total volume contributed by purchases of these households on a quarter to quarter basis;
3. The total number of brands and/or types introduced during these periods;
4. The total number of brands and/or types withdrawn during these periods;
5. The average number of items purchased per household by quarterly periods by brand and type;
6. The average volume purchased per household by quarterly periods by brand and type.

The data for the primary analysis in the study were tabulated by U.S. Total consisting of types of the product class shown within brands and also brands shown within types (flavors). Purchases made by over 5,000 households were examined. The first purchase of a new flavor could have been made in any of the four quarterly periods within a year. This quarter of first purchase was the base period in the analysis. Purchase behavior was examined in terms of prior and subsequent purchases to the base period. Data were evaluated on both a cumulative and non-cumulative basis in order to observe new and repeat buyer patterns as well as duplication among purchases of brand and types.

The average number of items purchased and the average volume per household has not changed significantly through time for the product under consideration. There have also been more types than brands introduced with very few withdrawals from the market of either type or brand. This phenomenon is an important factor as will be discussed later.

Research Design

The basic research design is quasi-experimental and is classified by Campbell and Stanley as having the general characteristics of a time series:

"The essence of the time-series design is the presence of a periodic measurement process on some group or individual and the introduction of an experimental change into this time series of measurements..... It can be diagramed thus following Campbell and Stanley (Campbell & Stanley, 1968):

$$O_1 O_2 O_3 O_4 X O_5 O_6 O_7 O_8"$$

With this design an initial examination was made of repeat buyers of the new flavor introduction who were previously loyal to the new flavor's brand. These data were then compared to repeat purchases of new buyers in order to closely determine the source of retention.

The second step in the analysis involved tracing the effects of the above buying patterns on the other types within the brand containing the new flavor as well as on competitive items. A Markov model was used for this portion of the analysis.

Product-Line Extensions and the
Product Life Cycle

Although client non-disclosure pacts limited this writer's investigation, he was able to obtain limited data from selected product classes other than the one used for the primary analysis of this study. Ten such products were made available for the purpose of observing changes in average volumes per household over a three-year period. Five of the product categories were classified as being in the growth phase of their product life cycle.³ Three of the products were purchased as frequently as once in every twenty days; four had purchase cycles of once in every thirty days; the balance had purchase cycles ranging from once in every forty-five to ninety days. Among all mature products, significant increases in purchase volumes were not observed among purchasing households. In addition, long-run market penetration rates and brand shares appeared to be in an asymptotic state among the five mature products. The year to year penetration rates (1967 through 1970), in terms of gains in household buying and buying rates, did not change significantly above population growth which averaged 1-1/2% per year for the three years of observation.

The five mature product classes all revealed total buyer penetration rates in excess of 50% with definite stable-state signs. The mature product class used for the primary analysis in this study has penetrated 80% of total U.S. households, a rate which held steady prior and subsequent to the introduction of new items during the three-year observation period.

The five product classes classified as being in the growth phase of their product life cycle have all been in existence ten years or less. Year to year trends for these products revealed statistically significant growth, in numbers of households buying and in purchase rates, above the average rate of population growth. In addition, these product classes have penetrated a range, thus far, of from only 20% to 50% of total U.S. households. Thus a relative range for continued market penetration for the newer products would appear to exceed that of the mature products.

The above discussion, however, does not change the meaning of the hypothesis in this study. If long-run growth is to be achieved then lasting volume must be obtained from outside the market, from competitors, or from light buyers. The new product brand cannot rely on its hard core heavy buyers to purchase additional volume. What is implied, however, is that with newer product categories, still in the early stages of growth, the hard core buying element has, in all likelihood, not had time to build to its ultimate level leaving room for expansion of lighter buyers into this category. In contrast with a mature product class, growth product classes may expect conversion of light and medium buyers into the heavy category until such time as the product moves into the leveling mature phase of its product life cycle.

The key point in the analysis of product-line extensions may well lie in the nature of the product class from which it comes, i.e., mature versus growth, and not just from the innovative nature of the line extension. It may, in fact, be possible to introduce two or three new flavors within a year in a brand that represents a growth product class with the result that lasting buyers are brought in at no expense to existing flavors, or at a

rate that offsets any proliferation between items. On the other hand, a mature product class that has achieved maximum penetration in the market to the extent that additional buyers are not significantly forthcoming, may find extensive proliferation upon existing items as line-extensions are added.

The crux of the whole matter is that the mature product has a more limited means for expansion than does the newer product. The newer growth product is still building penetration in the market as well as gaining volume from competitive brands. The mature product may be limited to only deriving its additional volume from competitors.

Table 1 relates share trends among flavors and brands of a growth phase product. The flavors shown are all within Brand H, and indicate the addition of two flavors to the line, one during period eight, and the other during period eleven. The product has been in existence for ten years.

Periods nine through twelve, as related in Table 1, produced a record volume for the product class. The primary reason for this rate of growth was the continuing acceptance of Brand H's products. The brand represents a strong innovation in its market, and, thus, far, competition has not competed effectively, either in quality or number and types of items available. The squeeze on competitive brands by Brand H, in terms of share, indicates not only gains from competition, but a general monopolization tendency. The brand has been able to bring new buyers into the market and, at the same time, convert competitor's buyers to its market. However, the existing loyal buyers to the brand did not increase their purchasing in total. Transaction size remained a steady 2.08 units throughout this period of growth for these heavy buying households. The brand was able to convert marginal buyers into heavy buyers, at a significant rate, thus expanding the buying base.

During periods one through seven, Brand H marketed eight items with a total brand share ranging from 57% to 59% of the total market. By the end of period twelve this share had increased to 75% of the market. Flavor I was introduced during period eight and immediately gained 18% of Brand H's total share, which increased during the period to 64%. By period nine Flavor I's share had increased to 39% of the total brand's share leveling finally to 30% by the end of period twelve. Following through the data in Table 1, it can be seen that Flavor I contributed substantially to the increased success of Brand H and cannibalized on only one flavor within the brand to any significant degree (Flavor F). The introduction of Flavor J in period eleven should add additional information to the growth pattern of Brand H through time. Additional data is needed, however, for any extensive analysis on this new introduction.

In general, Brand H has increased its share by a significant amount through time by converting non-buyers, competitive buyers, and light buyers to the more loyal hard core of its buyers. This was accomplished primarily through the effectiveness of one specific line extension, Flavor I.

Table 1

SHARE TRENDS AMONG FLAVORS AND BRANDS OF A
GROWTH PHASE PRODUCT

Brand/ Flavor	Periods											
	1	2	3	4	5	6	7	8	9	10	11	12
Brand A	.05	.05	.05	.07	.06	.06	.06	.05	.01	.01	.02	.01
Brand B	.11	.14	.08	.08	.06	.07	.07	.07	.06	.04	.05	.05
Brand C	.03	.03	.03	.03	.05	.02	.04	.02	.02	.02	.01	.02
Brand D	.08	.08	.12	.07	.09	.09	.09	.07	.04	.04	.06	.05
Brand E	-	-	-	.01	.01	.02	.01	.01	.01	.01	.02	.01
Brand F	-	-	-	.01	.03	.02	.02	.01	.01	.03	.01	.01
Brand G	.16	.14	.14	.16	.14	.16	.12	.13	.16	.12	.06	.10
Brand H	.57	.56	.58	.57	.56	.56	.59	.64	.69	.73	.78	.75
Flavor A	.05	.02	.02	.03	.03	.01	.01	.01	.01	.01	*	.01
Flavor B	.05	.05	.05	.04	.03	.03	.03	.03	.02	.04	.04	.03
Flavor C	.07	.08	.07	.05	.05	.07	.07	.07	.04	.07	.06	.06
Flavor D	.04	.05	.05	.07	.05	.05	.03	.03	.04	.03	.06	.03
Flavor E	.05	.06	.08	.06	.05	.04	.06	.03	.02	.04	.05	.03
Flavor F	.22	.22	.20	.25	.28	.28	.29	.25	.13	.17	.18	.18
Flavor G	.02	.03	.04	.03	.03	.02	.03	.01	.02	.01	.01	.01
Flavor H	.07	.05	.07	.04	.04	.06	.07	.03	.02	.01	.02	.01
Flavor I	-	-	-	-	-	-	-	.18	.39	.35	.29	.30
Flavor J	-	-	-	-	-	-	-	-	-	-	.07	.09

*Less than 1%

Analysis of Buyer Behavior Among Purchasers of
New Flavor of a Mature Product Class

Three analyses were conducted among the new and repeat buyers of the new flavor. The first examined total new and repeat buyers in order to determine the extent of general market acceptance. The second examined market acceptance of the new flavor among buyers of the new flavor brand. The third examined market acceptance of the new flavor among previous buyers of the new flavor brand. Tables 2, 3, and 4 reflect the results of these analyses.

In order to gain a true picture of the market success of the new flavor relative to the buyer groups identified above, it was necessary to convert the purchases to volume. During the first year, the average purchase size among all buyers was 20.5 ounces. These rates did not change significantly in the second year. Thus, this stable rate was applied to both first and second year purchases in order to obtain an estimate of total volume for all buyers. Table 2 reflects these estimates. Tables 3 and 4 reflect a breakdown of this total volume by new and previous buyers of the new flavor brand.

In general, the new flavor introduction achieved market acceptance. The data in Table 2 are based on total market penetration rates, among National Consumer Panel households, of 37% with an excellent first repeat rate of 54%. Second and third repeat rates were 33% and 20% respectively.

Table 2

New and Repeat Buyer Volumes for New Flavor Item by Total Buyers

<u>Buyers</u>	<u>1st Year</u>	<u>2nd Year</u>	<u>Total</u>
Total New Buyers	2,400	350	2,750
Frequency of Purchase	1.0	1.0	1.0
Avg. Transaction Size (ozs.)	20.5	20.5	20.5
Total Volume (lbs.)	<u>3,088</u>	<u>480</u>	<u>3,568</u>
Total Repeats	1,550	1,585	3,135
Frequency of Purchase	4.2	4.2	4.2
Avg. Transaction Size (ozs.)	20.5	20.5	20.5
Total Volume (lbs.)	<u>8,341</u>	<u>8,529</u>	<u>16,870</u>
Total Volume All Purchase (lbs.)	<u>11,429</u>	<u>9,009</u>	<u>20,438</u>

Source: The National Consumer Panel

Table 3

New and Repeat Buyer Volumes for New Flavor Item by New Buyers

<u>Buyers</u>	<u>1st Year</u>	<u>2nd Year</u>	<u>Total</u>
Total New Buyers	800	100	900
Frequency of Purchase	1.0	1.0	1.0
Avg. Transaction Size (ozs.)	12.0	12.0	12.0
Total Volume (lbs.)	<u>600</u>	<u>75</u>	<u>675</u>
Total Repeats	290	55	345
Frequency of Purchase	2.1	2.1	2.1
Avg. Transaction Size (ozs.)	12.0	12.0	12.0
Total Volume (lbs.)	<u>457</u>	<u>87</u>	<u>544</u>
Total Volume All Purchases (lbs.)	<u>1,057</u>	<u>162</u>	<u>1,219</u>

Source: The National Consumer Panel

Table 4

New and Repeat Buyer Volumes for New Flavor Item by Previous Brand Buyers

<u>Buyers</u>	<u>1st Year</u>	<u>2nd Year</u>	<u>Total</u>
Total New Buyers	1,600	250	1,850
Frequency of Purchase	1.0	1.0	1.0
Avg. Transaction Size (ozs.)	24.9	25.9	25.6
Total Volume (lbs.)	<u>2,488</u>	<u>405</u>	<u>2,893</u>
Total Repeats	1,260	1,530	2,790
Frequency of Purchase	4.8	4.6	4.6
Avg. Transaction Size (ozs.)	20.9	19.2	20.4
Total Volume (lbs.)	<u>7,884</u>	<u>8,442</u>	<u>16,326</u>
Total Volume All Purchases	<u>10,372</u>	<u>8,847</u>	<u>19,219</u>

Source: The National Consumer Panel

It must be concluded, however, that the new buyer segment of the new flavor market contributed relatively little in terms of total buyers for the product. The total new buyers shown in Table 3 represent 33% of the total buyer category. Yet, when calculations were made, its total volume represented only 6% of the total new flavor purchase volume. The primary problem was a general lack of repeat volume.

The dominant market for the new flavor was centered in the previous buyers of the new flavor brand. These buyers represented 67% of the total buyers and 94% of the total volume. Analysis of Table 4 suggests an initial stocking up among first purchases of the new flavor with a slight purchase decline among repeaters. Users of new and repeat buying models often seek such indicators in their search for future problem areas leading to market decline (Cunningham, 1961). In the case of this product, it can be expected that other items in the line may have suffered during the stocking up process with the new flavor. This will be discussed in the following section.

The Source and Disposition of New Flavor Buyers

The Markov Process

The Markov process used in this study assumes that the effect on a given state i at condition t is flowing from the probability distribution of all states in the system at only this condition, $t - 1$, and the transitional probabilities of moving from those states to state i during the conditional interval $t - 1$ to t . The assumption is that the knowledge of the probability distribution of this conditional period is sufficient to predict the probability distribution of states at the present and future conditional periods at the asymptotic state.

Two type of information are essential to carry forward the analysis. They are (1) an initial probability distribution (brand share), and (2) the transition matrix. The probabilities of the transition matrix are nothing more than the probability that a certain manufacturer will retain its customers. It is computed from a gain-loss summary. Given the transition matrix and brand shares for a particular data period, future brand shares can be predicted, i.e., probabilities, by multiplying the brand shares with the transition matrix raised to higher powers.

Analysis

Tables 5 through 7 illustrate flavor and brand loyalties of the new flavor buyers both prior and subsequent to their quarter of first purchase of the new flavor. The data in these tables were compared with predictions from the Markov model. The X^2 goodness of fit test was used to evaluate how well the model predicted actual brand and flavor shares prior and subsequent to the quarter of first purchase of the model to actual brand share data.

Table 5 reflects shares and predictions for five flavors and subsequent to the quarter of first purchase of the new flavor. The new flavor is an extension of Flavor A. The first impression upon reviewing the total flavor share information prior to the quarter of first purchase is one of stability of the shares period to period for the four quarters. It appears that the buyers may be in an asymptotic state.

To be in an asymptotic state should not be considered unusual given the maturity of the product class involved. The product has been in existence for decades and can definitely be classed in the maturity phase of the product life cycle.

The goodness of fit of the model reflects this stability. The transition matrix was fixed at $t - 1$ the fourth buying period prior to the quarter of first purchase. The predictions of the model were quite close to the actual observed shares in all periods. Using equivalent volumes for shares, at X^2 statistic was computed for all prior expected versus observed observations. The resulting 7.68 statistic did not exceed the .95 significance level at 19 degrees of freedom. The measures showed no significant differences.

The flavor share for the quarter of first purchase was not included in the predictions as it was considered an independent variable. The purpose in the analysis was to observe differences between expected and actual observations prior and subsequent to the insertion of the atypical variable. If the quarter of first purchase probabilities had been included in the predictions, the stability of the transition matrix would have been disturbed for subsequent periods. By using the quarter of first purchase as an independent variable it was possible to continue observing effects on the stable state. The greatest volatility on the disposition of buyers side of Table 5 appeared in those quarters adjacent or near adjacent to the quarter of first purchase. In terms of total periods, however, it can be observed that flavor shares fell near to ones observed in prior periods. The X^2 test for goodness of fit in the subsequent periods reflected no significant differences between predicted and actual observations. Thus, in terms of the complete set of data the new flavor introduction caused little long-run volatility between flavors.

Table 6 reflects shares and predictions for seven brands prior and subsequent to the quarter of first purchase of the new flavor. Again, a relatively stable state can be observed between quarter to quarter purchase shares in the prior period. X^2 for the total prior period was computed at 14.65 which was below the .95 significance level with 27 degrees of freedom.

Brand A contains the new flavor. It can be observed that from the fourth period prior to the fourth period subsequent to the quarter of first purchase, Brand A increased its market share by only one point. In terms of share, Brand A was unable to hold long run purchase volume equivalent to the volume generated in the quarter of first purchase. An examination of competitive brands will reveal a converse condition to Brand A in most instances.

In general, all competitive brands to Brand A, with the exception of Brand F, suffered share losses during the quarter of first purchase of the new flavor. All recovered, however, to levels approximating brand shares generated in the prior periods. The volatility that was generated by the new flavor introduction did cause a significant difference between the model's predictions and observed values during the subsequent period. X^2 was computed at 75.32 indicating significance at the .999 level. The values that contributed primarily to the high statistic occurred in the first subsequent period to the period of first purchase. Differences in expected and observed shares between Brands B, D, and F, contributed heavily to these differences. The implication, again, is that short-run and not long-run changes were prevalent.

Table 5

Index of Brand Flavor Loyalty Prior and Subsequent To
First Purchase of New Flavor

	1	2	3	4	X ^a	5	6	7	8
Flavor A	.42 (.41)	.41 (.42)	.42 (.42)	.44 (.43)	.67	.47 (.43)	.47 (.44)	.45 (.44)	.43 (.44)
Flavor B	.39 (.38)	.42 (.39)	.40 (.40)	.41 (.40)	.21	.39 (.41)	.39 (.41)	.41 (.42)	.42 (.41)
Flavor C	.10 (.09)	.08 (.09)	.09 (.08)	.08 (.08)	.06	.06 (.08)	.06 (.07)	.06 (.07)	.08 (.07)
Flavor D	.04 (.06)	.06 (.05)	.05 (.06)	.04 (.05)	.03	.03 (.05)	.03 (.04)	.04 (.05)	.04 (.04)
Flavor E	.05 (.06)	.03 (.05)	.04 (.04)	.03 (.04)	.03	.05 (.03)	.05 (.04)	.04 (.04)	.03 (.04)

a Denotes quarter of first purchase of new flavor. New flavor is a variation of Flavor A and an addition to Brand A. First purchase could have been made in December 1967-February 1968, March-May 1968, June-August 1968, or September-November 1968. () indicates computerized predictions.

Table 6

Index of Brand Loyalty Prior and Subsequent to First Purchase of New Flavor

	1	2	3	4	X ^a	5	6	7	8
Brand A	.20 (.23)	.23 (.21)	.21 (.22)	.21 (.22)	.53	.22 (.22)	.23 (.22)	.21 (.22)	.21 (.22)
Brand B	.33 (.30)	.31 (.32)	.32 (.32)	.32 (.32)	.20	.26 (.32)	.28 (.30)	.29 (.29)	.31 (.29)
Brand C	.07 (.08)	.05 (.07)	.06 (.06)	.06 (.06)	.04	.08 (.06)	.08 (.07)	.08 (.08)	.06 (.08)
Brand D	.09 (.08)	.10 (.09)	.10 (.09)	.09 (.09)	.05	.05 (.09)	.05 (.07)	.04 (.05)	.05 (.05)
Brand E	.10 (.11)	.11 (.11)	.09 (.10)	.10 (.10)	.06	.10 (.10)	.09 (.10)	.10 (.10)	.10 (.10)
Brand F	.10 (.11)	.08 (.10)	.07 (.09)	.06 (.07)	.06	.15 (.07)	.13 (.10)	.15 (.12)	.12 (.13)
Brand G	.11 (.09)	.12 (.10)	.15 (.12)	.16 (.14)	.09	.14 (.14)	.14 (.14)	.13 (.14)	.15 (.13)

a Denotes quarter of first purchase of new flavor. New flavor is a variation of Flavor A an addition to Brand A. First purchase could have been made in December 1967-February 1968, March-May 1968, June-August 1968, or September-November 1968. () indicates computerized predictions.

Table 7

Index of Brand Flavor Loyalty Prior and Subsequent to First
Purchase of New Flavor Within Flavor A Purchases Only

	1	2	3	4	X ^a	5	6	7	8
Brand A	.30 (.31)	.34 (.31)	.32 (.32)	.29 (.32)	.72	.37 (.31)	.37 (.33)	.34 (.34)	.30 (.33)
Brand B	.20 (.19)	.19 (.19)	.21 (.19)	.23 (.20)	.08	.15 (.21)	.15 (.19)	.16 (.17)	.17 (.17)
Brand C	.07 (.07)	.05 (.07)	.04 (.06)	.04 (.05)	.01	.04 (.05)	.04 (.04)	.04 (.04)	.03 (.04)
Brand D	.14 (.12)	.10 (.12)	.09 (.11)	.10 (.10)	.01	.03 (.10)	.04 (.07)	.06 (.05)	.10 (.06)
Brand E	.05 (.07)	.07 (.06)	.06 (.06)	.08 (.06)	.07	.07 (.07)	.07 (.07)	.09 (.08)	.05 (.08)
Brand F	.09 (.11)	.10 (.11)	.12 (.11)	.10 (.11)	.07	.16 (.10)	.15 (.13)	.15 (.14)	.17 (.15)
Brand G	.15 (.13)	.15 (.14)	.16 (.15)	.16 (.16)	.09	.18 (.16)	.19 (.17)	.15 (.18)	.18 (.17)

a Denotes quarter of first purchase of new flavor. New flavor is an addition to Flavor A within Brand A. First purchase could have been made in December 1967 - February 1968, March-May 1968, June-August 1968, or September-November 1968. () indicates computerized predictions.

Table 7 reflects shares and predictions for seven brands, prior and subsequent to the quarter of first purchase of the new flavor, for total Flavor A purchases only. This view affords a direct comparison of the effects of Brand A's flavor extension relative to similar flavors within brands.

Although more variance between expected and observed shares can be observed in this table than in the previous two, a relatively good fit still existed in the prior observations. X^2 was computed at 25.80, less than the .95 level of significance with 27 degrees of freedom. Yet, this statistic was the highest recorded for the prior observations.

A striking point in Table 7 is the fact that Brand A's share of Flavor A fell to its original prior brand share during the fourth quarter subsequent to the first purchase of the new flavor. It should also be recalled that Brand A gained an overall one share point (Table 6) by the end of the fourth subsequent period. This indicates that in terms of Flavors A's family of products the overall long-run effect for Brand A reflected no change. Yet the activity generated by the new flavor buyers.

Brand A did, however, maintain an above-average share of Flavor A for three periods subsequent to the first quarter of purchase. During these periods most competitive brands retained a below average market share. The X^2 value for subsequent periods was the highest observed among the subsequent purchase categories. A 90.86 statistic was computed indicating significance at the .999 level for 27 degrees of freedom.

Conclusions

The hypothesis of this study suggested that long-run purchase volume of households loyal to a particular brand will not be affected by purchases of a new type introduced within that brand. It was discovered that the bulk of volume for the flavor did indeed come from previous loyal buyers of Brand A. This finding, coupled with the fact that Brand A was unable to hold, in the long-run, purchase volume generated by the new flavor introduction, reflected the general behavior pattern suggested in the hypothesis. The indication is that the loyal buyers did not increase their total consumption of the product in the long-run market growth for Brand A was not achieved.

A buyer may experience genuine discontent with a product implying the need for an extended search process for new products. Or, the need for temporary or intervening purchases may arise to relieve the boredom of routine related to content purchases of the same products. After one or two intervening purchases this type of buyer will normally return to previous purchase patterns. The inference from the data in this study is that, in general, purchasers of the new flavor behaved in this latter manner. An examination of the inferences resulting from the data should lend additional insight into why this conclusion has been reached.

A general stability of period to period brand shares among households was evident from the source and disposition analysis. This apparent asymptotic state was not considered unusual given the maturity of the product class and its buyers. The basis assumption for the application of the Markov model, i.e., that learning had occurred was, thus, fulfilled.

In the early periods of a product class, the gains and losses from brand to brand can be of high magnitude. But as buyers travel up the slope of the learning curve to its asymptote (equilibrium), gains and losses between brands become smaller and smaller until just before the asymptote they are infinitesimally small. This was the condition noted in the source and disposition analysis.

The asymptotic state of equilibrium may also reflect a leveling out of buyer penetration. In other words, a mature product class in a state of equilibrium relative to the matrix of transition probabilities between brand shares may also be in a state of equilibrium concerning the number of new buyers entering the market. With the same buyers purchasing the product from period to period, it should not be surprising that a rather static long-run market situation can develop given a relatively low level of product innovation.

Different decision strategies vary with respect to the position of the consumer on the product learning curve. The pure strategy of invariably choosing the previously purchased item may be expected to yield less disturbance in the consumer's overall routine. With respect to variability of behavior, mixed strategy of selecting different brands and/or flavors on different occasions may actually reinforce loyalty for the more frequently purchased item by reducing overall boredom for the product class in general.

The new flavor buyers were willing to give up income in the short-run by purchasing the new flavor in addition to other items within the same line. This behavior was not retained by Brand A in the long-run, however.

The buyers of the new flavor item originated primarily from the overall group of loyal buyers of the new flavor brand. The inference is that these buyers had reached an upper limit of product that they are willing to buy. Thus, in terms of building market share and volume, Brand A was limited to deriving long-run volume from outside the current market and from competitors who might also contain light buyers of Brand A's products. Since an equilibrium ceiling had been reached in the overall product category with reference to buyer penetration (penetration currently exceeds 80% of total U.S. households), the primary means of expansion for Brand A, via the new flavor introduction, was limited to competitive sources and occasional buyers. The findings point to a definite failure on the part of Brand A to exploit its one source of long-run volume. For example:

1. The new flavor introduction was not responsible for long-run volatility between the various flavors.
2. Brand A was unable to hold long-run purchase volume anywhere near the volume generated in the quarters of first purchase of the new flavor.
3. Conversely, competitive brands suffered only short-run losses in volume and share because of the introduction of the new flavor, and enjoyed a return to previous higher levels by the end of the fourth subsequent quarter after the periods of first purchase of the new flavor.
4. Brand A's share of Flavor A fell to its original prior level by the end of the fourth subsequent purchase period, although Brand A maintained an above-average share for this flavor for three of the subsequent quarters.

5. During the periods of first purchase of the new flavor, Brand A's other flavor items suffered share losses indicating proliferation from the new flavor item. Gradually, however, this share has been increasing in subsequent periods and is nearing prior levels.
6. A general pattern of substitution between various Brand A flavors was noted as subsequent volume totals returned to prior levels.

These findings reflect definite signs of buying households engaging in the "psychology of complication." In the words of Howard and Sheth:

A surprising phenomenon, we believe, occurs in many instances of frequently purchased products, such as grocery and personal care items. The buyer, after attaining routinization of his decision process, may find himself in too simple a situation. He is likely to feel monotony or boredom associated with such repetitive decision making. It is also very likely that he is satiated with even the most preferred brand. In both cases, he may consider all existing alternatives including the preferred brand to be unacceptable. He, therefore, feels a need to complicate his buying situation by considering new brands, and this process can be called the psychology of complication. The new situation causes him to search for identity with a new brand, and so he begins again to simplify in the manner described earlier. Thus with a frequently purchased item, buying is a continuing process with its up and downs in terms of information seeking, analogous to the familiar cyclical fluctuations in economic activity (Howard & Sheth, 1969).

In general, the findings indicate a group of buyers unwilling to purchase additional product in the long-run, but willing to vary their behavior in the short-run, perhaps because of the reasons noted by Howard and Sheth. Identification of where the buyer is located relative to the "simplification-complication" process in terms of information seeking is clearly very important to the marketing manager. If it can be determined, for example, that a group of buyers are at a level of routinization where satiation or monotony with regular brand buying is apparent, then a new brand or flavor could possibly be introduced that would provide the needed source of change that will ultimately build a greater share of market for the innovator. What is difficult to determine, however, is the extent of the discontent with existing routinization. If, as in the case of the product under review in this study, it is a short-run phenomenon that does not have the strength to hold switching buyers from other brands then long-run effects will more than likely be negligible in terms of increased overall sales of product. A true innovation, on the other hand, could satisfy the need for new product searching, and thus, become a vehicle toward increased market growth.

Footnotes

1. J. Taylor Sims is Assistant Professor of Marketing, University of South Carolina.
2. Of course, the strategy involving product-line extensions can also be of a defensive nature, e.g., the matching of competitors' products in order to hold one's current market position. The assumptions in this analysis are based on the firm's inherent desire for market growth and its ability to lead from strength.

3. Data furnished by the Market Research Corporation of America. Product categories and actual counts cannot be disclosed. Definitionally, a product is considered, in this study, to be in the mature phase of its product life cycle if an asymptotic state has been reached in terms of new buyers entering the product class, i.e., increased new buyer penetration for the entire product class is not exceeding the average rate of population growth per year.

References

- Brown, G. Brand Loyalty--Fact or Fiction. Advertising Age, 1962, 23, 53-55.
- Campbell, D. T. & Stanley, J. C. Experimental Designs for Research. Chicago: Rand, McNally, 1968.
- Cunningham, R. Customer Loyalty to Store and Brand. Harvard Business Review, 1961, 39, 127-137.
- Fourt, L. A. & Woodlock, J. W. Early Prediction of Market Success For a New Grocery Product. Journal of Marketing, 1960, 25, 31-38.
- Frank, R. E. Brand Choice as a Probability Process. Journal of Business, 1962, 35, 35-56.
- Howard, J. W. & Sheth, J. N. The Theory of Buyer Behavior. New York: John Wiley & Sons, 1969, 27-28.
- Howard, R. A. Dynamic Inference. Operations Research, 1965, 13, 35-42.
- Kuehn, A. Consumer Brand Choice as a Learning Process. Journal of Advertising Research, 1962, 2, 10-17.
- Massy, W. F. A Stochastic Evolutionary Adoption Model for Evaluating New Products. Working Paper No. 95, Graduate School of Business, Stanford University, 1967.
- Market Research Corporation of America. The National Consumer Panel. Unpublished Monograph, 1970.
- Parfitt, J. H. & Collins, B. J. K. The User of Consumer Panels for Brand-Share Predictions. Journal of Marketing, 1968, 5, 131-145.

CONCEPTUALIZING ADEQUACY OF INFORMATION*

John A. Howard
Columbia University

Introduction

In the last decade, our institutions and ideas have come under severe fire by those who are attempting to formulate a better concept of the good life for society. An element of that concept has come to be adequate information for the consumer. Awareness of this requirement is being reflected in executive, judicial, legislative and regulative bodies at local, state and national levels.

How to conceptualize the idea of adequate information so as to develop criteria for action directed to insuring that the consumer does have adequate information, in fact, is not exactly clear. Even less clear are the concepts and techniques required for obtaining data appropriate to whatever criteria are formulated. Permit me to attempt an answer to this two-fold problem.

I propose four criteria for deciding whether information to the consumer is adequate. For these, I am indebted to the staff of the Bureau of Consumer Protection of the Federal Trade Commission. Information to be adequate should have the following characteristics:

Truthful
Intelligible¹
Relevant
Complete

My purpose is to conceptualize these criteria in such a way that theory and empirical research can be brought to bear in providing data for applying them quantitatively in concrete instances. In this way, the proposed structure can be tested for its correspondence with reality and the scientific legitimacy of the criteria established.

The first two criteria, however, will be neglected. General agreement exists among most all interested parties--consumers, industry and the government--that consumer information, including advertising, should be truthful. Differences of opinion arise in specific cases, but the general principle is widely accepted. For information to be intelligible would seem to be very much in the interest of the dispenser of the information, the advertiser, for example. Effort is devoted to achieving this in advertising agencies and their client organizations. Consequently, it should largely take care of itself. Both characteristics are important and deserving of research, but they lack the urgency of the last two, relevancy and completeness.

Relevancy

To ask what information is relevant is to ask how the consumer conceptualizes the object which the information is about. Thus our task is to describe how the consumer conceptualizes the brand.

Research, especially that in concept formation, indicates that he conceptualizes the brand in denotative, descriptive or non-evaluative dimensions as well as evaluatively. We label a summary measure of these denotative dimensions

"brand comprehension." He uses them to identify the brand and to discuss it with his friends and neighbors. These dimensions are not numerous, being severely limited by his capacity to process information. Also, he has less difficulty conceptualizing these denotative dimensions for a concrete object than for an abstract thing such as a service. Thus, we can think of brand comprehension as an aspect of a brand image or concept.

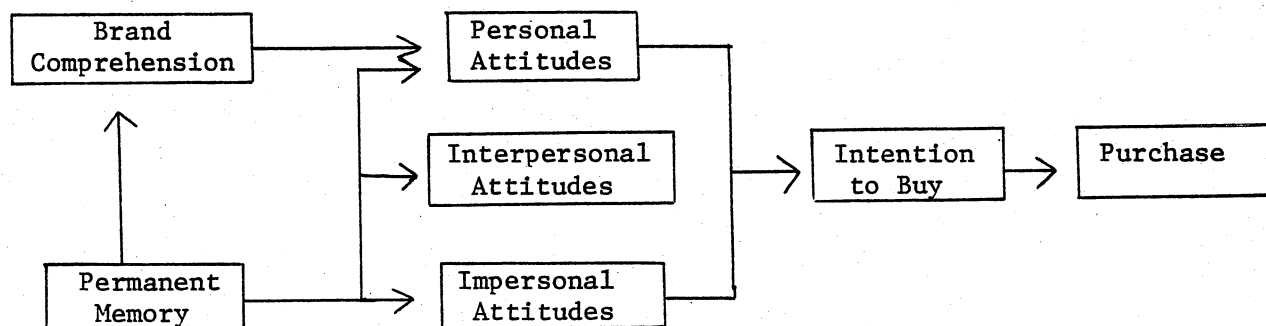
Another aspect of a brand concept is the evaluative dimensions. What characteristics of the brand cause him to say it is good or bad for him in terms of his motives? The extensive literature in attitude research is quite helpful here. Unfortunately, it has two limitations. First, with increasing exception, it deals with unidimensional attitudes, whereas our concern is with multidimensional phenomena. Second, the research in attitude, has with few exceptions, not been linked to the denotative or brand comprehension type of research characterizing concept formation.

It is important that we distinguish among two classes of attitudes: personal and interpersonal. An ad tells you, "Brand X is sweet and tastes good; it will make you beautiful and attractive to others." We can call the first "personal" and the second "interpersonal." Interpersonal or self-concept appeal is an assertion or implication that the buyer can change, in a favorable way, how he views himself or how others view him. Consequently, the distinguishing feature of the self-concept appeal is "others," and it has a two-fold aspect. First, it is one of the most seriously criticized features of modern advertising. Second, it is probably more difficult to substantiate than the personal dimensions.

It is the combination of denotative and connotative dimensions that constitute a brand concept. Information is relevant whenever it helps the buyer place the brand on the dimensions of this concept.

Finally, our discussion of attitudes above implicitly dealt with either personal or interpersonally, things that matter to him directly and immediately. Still another set of conditions prevail and these are particularly associated with the act of purchase. Price of the product is an obvious example. The buyer does not care about the price per se, but he does care about how much of other things he must forego to buy it. Also, his expectations about whether the brand is available at his regular retailer is a dimension. Finally, his attitude toward how people with power over him will feel about his purchasing the brand also matters. Because these dimensions do not matter personally or interpersonally, but instead are looked upon as constraints, they are called impersonal attitudes.

In summary, any information which enables the buyer to place a brand on each of these four sets of dimensions--brand comprehension, personal attitudes, interpersonal attitudes and impersonal attitudes--is relevant. These sets of dimensions can be thought of as being related in the following way.

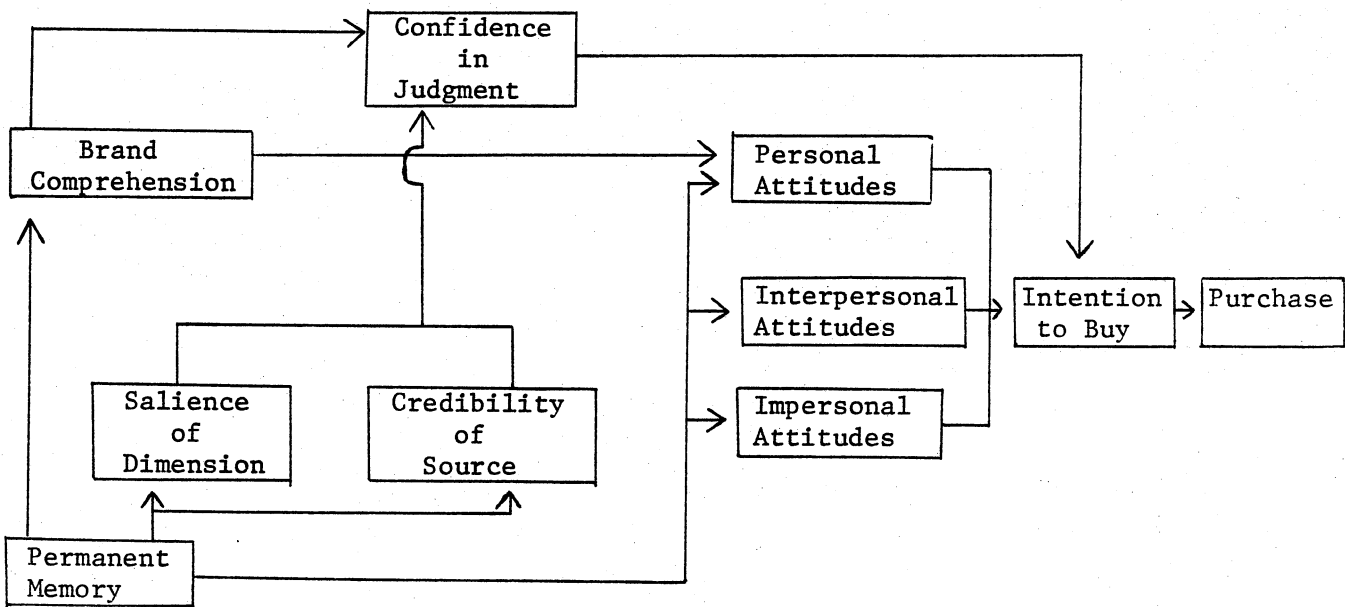


We must add a construct representing information--permanent memory--in developing a testable structure. We do not discuss the mechanisms that regulate the input to the buyer's memory: How did the information get there in the first place? The important point here is that it provides the informational input to the other constructs.

As implied by the connecting arrow, brand comprehension forms an essential foundation for personal attitudes to develop. Personal, interpersonal and impersonal attitudes jointly cause intention to buy. Intention to buy is incorporated because purchase is a discrete, infrequent event and to use it as the independent construct places a burden on research. Intention measures, on the other hand, can be taken at any time. Intention, in turn, causes purchase.

Completeness

When does the consumer have enough information? Psychologically we think of this as occurring when tension engendered by the conflict arising from uncertainty is at some comfortable level. A construct that describes this psychological state is confidence in judging the quality of the brand. We postulate that when the consumer has received adequate information on the four sets of dimensions, his confidence will be high, and we can conclude that in a vitally subjective sense, his information is complete. The following structure is postulated.



Brand comprehension can be clearly thought of as both a specific content construct and an amount of information construct. In the first capacity, it shows how well he knows the criterial, denotative dimensions and so serves as a basis for evaluative feelings to develop. In the second capacity, it seems to us that he has a subjective sense of how accurately these judgments conform to reality. The human has survived, in part, because he has been rewarded when he judged accurately and punished when he judged inaccurately.

All three types of attitudes pose a problem, however. Their evaluative nature complicates formulating their role in completeness of information. To

the extent that the consumer has information about where the brand lies on each set of dimensions should per se have no effect on confidence. This would indicate merely his evaluation, not how confident he is in that evaluation. Consequently, we must look elsewhere for the relation of attitudes and completeness and so to confidence.

The credibility of source of evaluative information is a factor. To the extent the buyer believes the source in terms of both its expertise and its motives, he will be affected by it.

Also, presumably some of the evaluative dimensions are more important or salient to the consumer than others; some relate to motives that are higher in his motive hierarchy than others. This is clearly borne out in data on personal attitude dimensions. If so, information on these crucial ones should provide a relatively greater boost to confidence than information on less important dimensions. True, a growing body of evidence suggests that this importance weighting does not influence behavior directly, but I will postulate that it exerts an indirect effect by guiding the consumer's selection of information. He selects that which applies to the more salient dimensions in preference to information about the less salient dimensions.

Thus, quantity of information is basically the number of relevant descriptive and evaluative dimensions--personal, interpersonal and impersonal--on which he has received information. Thus, we conclude that confidence is a function of both brand comprehension and evaluative information taken into long-term memory, weighted by its credibility and the salience of the dimension to which it pertains.

Finally, it is probably obvious that a single ad could not be expected to completely inform a buyer of a radically new product, for example, where new choice criteria must be formed. Nevertheless, it can be judged on whether it contributes to completeness.

Implications

Formulating these two criteria--relevance and completeness--in terms of an explicit structure that can be modelled empirically is essential because of complex policy implications. Let us illustrate by referring to television.

Industry experience is apparently clear that when using a recall criterion, the 30-second commercial is economically superior to the 60-second. Further, some evidence indicates that even with a sales criterion, the 30-second is still economically superior to the 60-second, except in the case of new products. Thus, for industry generally to provide more information than it now does, we tentatively conclude, would be uneconomic in its own terms.

Two possibilities arise. First, we can resolve the issue by concluding that it is the obligation of industry in fulfillment of its privilege of serving the consumer, to provide the additional information made possible by a 60- or 90-second commercial. Better evidence than we now have that the consumer really needs more information than can be provided in a 30-second commercial would have to be generated. Hence, the theoretical system set forth here should be modelled to determine precisely how much information is needed to insure that it does exceed that possible in a 30-second message. This would be a cost-benefit approach.

But second, perhaps the content of the 60-second ad should be changed, instead of extending it to a longer commercial. For this we must know how the 30-second commercial is achieving its effect. Is its role heavily motivational and less cognitive, for example? Perhaps by redesigning the content of the 30-second commercial it will provide enough information. To answer these questions we must apply the relevancy and completeness criteria by running an ad, use the structure and examine the content of the ad, the level of the variables and strength of the linkages among the constructs.

Footnotes

1. The bureau terminology is "comprehensible" but it can easily be confused with the label for a construct used here. I have renamed it.

CONSUMER PROTECTION AND
THE VALUE OF INFORMATION

Robert A. Mittelstaedt
University of Missouri

Th' enthusyasm iv this counthry, Hinnissy,
always makes me think iv a bonfire on an icefloe.
It burns bright so long as ye feed it, an' it
looks good, but it don't take hold, somehow, on
th' ice. ---Mr. Dooley

Introduction

The central thesis of this paper is a simple one based on several assumptions: (1) If "consumerism" is to continue as a viable force in public affairs, programs to protect the consumer must enjoy a widespread level of public support. (2) To insure public support consumer protection programs must provide noticeable and substantial benefits to the consumer. (3) The provision of useful information to the consumer is an important form of consumer protection.

There may be debate about any of these three points. One could argue that public support is not necessary or that, in any event, it is independent of the results produced by various programs. Further, one could argue that information programs should rank considerably below product safety programs or anti-trust enforcement as vehicles for protecting the interests of consumers. However, this paper will take these three assertions as "given" and attempt to develop the idea that the concept of the "value of information" can be a useful guide to the construction of priorities for both those agencies charged with protecting the consumer and those researchers who wish to make their efforts relevant to consumer problems.

The value of information is a function of the expected opportunity loss (probability and cost) of suboptimal choices. If we accept the idea that "optimality" exists in the eye of the beholder, the ultimate judge of the value of information is the consumer. To measure this value, and thereby set our priorities, we must look at the consumer, his perceptions and his behavior. However, this paper is not simply a call for "more research" or an admonition that regulatory agencies should remain inactive until "more is known." There are pressing problems which demand attention now. Since resources for agencies and researchers are limited, some initial agreement on priorities will have to be based on a priori estimates of costs and probabilities. Because this organization would seem to be an ideal forum for "protectors" and "researchers" to explore common ground for establishing mutual priorities, this paper is an attempt to spark a dialogue. To this end the paper will look at several information programs, attempt to assess their probable value to the consumer and examine the implications of this estimated value for protection agencies and researchers.

Beef Grading

First, consider a very old information program--grade labeling of beef. Unlike the mandatory inspection program which concerns itself with the basic

wholesomeness of the product, the grade labeling standards provide a means of describing product quality in uniform language--the familiar grades of "U.S. Prime," "U.S. Choice," "U.S. Good," etc. Because the grading program is voluntary not all beef is graded. However, the U.S.D.A. estimates that 80-85 percent of all fresh beef bound for the household market is graded [U.S.D.A. 1969, p. 102; Hutchinson, 1970, p. 1]. Of all beef graded in 1970, 80 percent was graded "U.S. Choice" [U.S.D.A. 1971, p. 147]. This figure, when considered with the observation that many retail outlets adhere to a "one-grade policy," suggests that the U.S. housewife is very likely to select "U.S. Choice" beef from her retail meat counter because she has no alternatives.

The concept of the value of information tells us, then, that as the probability of the typical housewife choosing "U.S. Choice" approaches 1.0 (because of lack of alternatives), the value of the information represented by the grade label to her approaches 0.0. Viewed in this light, it does not seem surprising that most housewives know very little about the grading system used, as a recent study documented [Hutchinson, 1970]. This does not mean that grading should be abandoned, for it has other useful purposes. However it does suggest that putting resources into consumer education programs about beef grading is not likely to achieve much until the grading scheme is revised to provide grade labels which are more "information-laden." Behavioral research could aid in the construction of a grading system which allowed discrimination on dimensions of significance to the housewife.

Unit Pricing

Unit pricing is another example of a program designed to provide useful information to consumers. Studies have shown that consumers find it difficult in the absence of unit price information, if not impossible, to find the "least cost alternative," even when instructed to do so [Friedman, 1966; Houston, 1972]. Further, these studies have shown that, compared to the "best buys," the actual choices consumers make "cost" them. It has been concluded that the consumer could save money if unit price information was made available.

The operative word in the previous sentence is "could." By contrast, the value of information concept directs attention to the use people make of unit-price information and would measure the value on the basis of behavioral change. For example, looking at one table from the "Kroger study," there appears to be a shift of higher income consumers toward lower unit priced alternatives in cooking oil [McCullough & Padberg, 1971]. If the goal of unit pricing is to facilitate this shift, the value of the unit price information is measured by the magnitude of the shift in buying patterns. In these data, the average net savings is about .08 cents on every pint sold.

In their excellent review of the Kroger study and others, Monroe and LaPlaca have shown that there are serious methodological pitfalls in studies of this kind, which is to say that they may not all be "fair tests" of the program's value [Monroe & LaPlaca, 1972]. Furthermore, it is clear that consumers may use unit-price information in other ways--to reinforce previous choices or to justify shifting to higher unit priced alternatives. Elsewhere Monroe has suggested a model for dealing with this expanded concept of information [Monroe, 1971].

Initially, research effort should be directed toward the design of methodologically sound experiments, the development of adequate educational

Table 1
 Cumulative Distributions of Sales of Cooking Oil,
 Distribution Deviations for Test and Control,
 High Income Stores, Period 4, Spring 1970¹

Unit Price Per Pint	Test	Control	Difference
Cents		Percent	
28.62	6.7	0.7	+6.0
29.33	13.9	6.5	+7.4
31.12	18.9	12.4	+6.5
31.33	24.6	19.6	+5.0
32.44	29.8	23.3	+6.5
33.00	38.8	38.2	+0.6
33.00	44.5	47.7	-1.9
33.26	56.5	60.0	-3.5
34.10	62.5	65.3	-2.8
35.33	71.7	72.9	-1.2
36.50	75.7	76.5	-0.8
36.66	84.5	88.7	-4.2
38.50	94.2	95.3	-1.1
41.11	100.0	100.0	0.0

¹Part of Table 8 [McCullough & Padberg, 1971].

programs about unit pricing and the further development of a model to assess the full range of uses of the information. Ultimately, the value of unit price information rests, not on how much consumers could save, but on how, and to what extent, they use unit price information in their buying decisions. When finally computed, the value should be compared to the costs of providing the information. Admittedly, it will not be an easy task but it appears to be, in principle, a solvable problem. Finally, assuming the research is "sound," the advocates of unit pricing should be prepared to accept the results.

Deceptive Advertising

At first glance, the concept of the value of information might not appear to be relevant to the area of deceptive advertising. After all, no one can argue that false and misleading advertising is of value to the consumer. Furthermore, legally removing deceptive advertising claims from the market place does not assure that the remaining (presumably) "true" claims will provide any useful information to the consumer. However, if our initial assumption about continued public support was sound, we should direct

our limited regulatory resources toward those deceptive ads which represent the greatest potential disutility to the consumer. The concept of the value of information, then, should turn us away from priorities based on the magnitude of the falsehood of particular claims and toward a set of priorities based on the probability that consumers may be misled and the costs, to them, of the deception.

This is not to say that the successful prosecution of a deceptive advertising case should rest on evidence that people were, in fact, misled, for the courts have already decided otherwise [FTC, 1944]. Nor does it imply that we spend endless hours debating what consumers' optimum and suboptimum choices might be. However, the concept of the value of information implies that the costs of deception are a function of the probability of being misled and the associated opportunity loss.

In constructing a priority scheme based on the likely costs of deception, first priority should go to insuring the truthfulness of claims about products involving physical risk. Even when the probability of being misled is small, the costs of following deceptive claims about the safety or healthfulness of some products is high.

When considering potential economic (as opposed to physical) disutility, the "costs" become less clear. The dollar amount of the purchase may be some guide but, without some probability estimate, it is not a perfect criterion. One might argue that the shoddiest selling practices tend to be associated with those relatively expensive goods and services that are purchased very infrequently--the "once-in-a-lifetime" purchases for most people--and attention might be directed to them first.

However, most advertising involves relatively low-prices, frequently purchased products. Calls for truth-in-advertising usually argue that the consumer is exposed to an overwhelming number of ads for products of this type and is unable to evaluate the claims made for them. This is undoubtedly true, but it does not follow that all of these claims, even if relied upon, represent equal potential losses to the consumer. With repetitively purchased products, the greatest disutility to the consumer comes when the consumer is unable to evaluate the claims made for the product, even after purchase, and continues to buy the product because he cannot refute the deceptive claims based on his own experience.

After all, there are some ad claims, the truth of which the consumer is able to judge for himself after purchase. To take several examples from ad claims recently "under fire," it would appear that the consumer is reasonably well-equipped to judge whether or not: (1) a crunchy candy bar crunches as loudly as it was made to appear to do in an ad, (2) a chain store has stocked the goods it has advertised, or (3) a soup contains as many vegetables as it was pictured to have.

On the other hand, the consumer is very poorly equipped to judge whether or not: (1) a disinfectant significantly reduces the incidence of common diseases, (2) a dentifrice reduces cavities, (3) a fruit drink contains a particular vitamin, or (4) a gasoline reduces air pollution.

Again, this is not to argue that false and misleading advertising should be allowed to flourish, but only that the "cost of being misled" can be

determined, in part, by looking at the question of whether or not the consumer is able to judge ad claims for himself, after the purchase, and alter his behavior accordingly. Researchers interested in making a contribution in the area of deceptive advertising might proceed by discovering the salient dimensions of products, for "false" claims may or may not be "relevant."

The Relevance of Research

As a final point, consideration should be given to the development of a mutual understanding of the research process and what it can, and cannot, contribute. At the risk of muddying the water with a study which is at the core of a controversial issue, consider the "Johns Hopkins' aspirin study." Although a "clinical study," it is not fundamentally different from the research conducted by most of the people in this room.

De Kornfeld, Lasgna and Frazier [1962] administered 5 popular aspirin-containing analgesics and a placebo to 198 post-partum patients who had not received a general anesthetic nor had suffered obstretic complications. Upon patient request, a randomly selected treatment was given each patient and her responses to a question which asked her to categorize her pain as "absent," "slight," "moderate," "severe" or "very severe" before the administration of the dosage and at 15, 30, 45, 60, 120, 180 and 240 minute intervals thereafter were recorded. A decrease in pain from the preceding level to the next level was counted as a score of "+1," a drop of 2 levels was counted as a score of "+2," and so forth with zero's (no change) and negative scores possible. Mean "pain-relief scores" were calculated for each treatment. The major conclusions of the study, which rest on the differences in these scores as tested for significance by Harter's modification of Duncan's multiple range test, were that "all five preparations gave better relief from pain than did the placebo and there were no significant differences among them during the first 3 hours of observation" [De Kornfeld, et. al., 1962, p. 75].

In assessing the results of this study and its ability to support these conclusions, the researchers among us may wonder if the random assignment of treatments overcomes the problems associated with the sample and the crudeness of the measurement instrument. The statistically-minded may question the conclusion that all 5 compounds are equally effective which is based on the acceptance of a null hypothesis and wonder, therefore, how one might prevent the demonstration of "significant differences in effectiveness" from being something other than a simple function of sample size. The public policy maker might wrestle with the question of whether the inter-individual variability (assuming that it is not just "noise") is sufficient justification for preserving wide consumer choice.

But if we all step back one pace and look at this study and its conclusions from the viewpoint of the consumer, we will probably question the value of information about "comparative effectiveness" which, when finally revealed, consists of the perceptions of effectiveness by a sample of consumers. In short, this "clinical study" provides data no different in kind from that which any consumer could generate for himself at modest cost with little risk. In fact, from the consumer's point of view, given inter-individual variability, the data from this study, based as it is on "averages," will fail to square with many consumers' experiential data. When that happens, it is not difficult to predict which data will be disbelieved.

Summary

Again, it should be stated that it was not the purpose of this paper to make an apology for current marketing practices nor to dismiss consumer information programs as useless. There are urgent problems which demand our attention. What is needed is some shared sense of purpose among the public policy makers and enforces and those of us, who by training and disposition, are equipped to study how consumers behave. Ultimately, if consumer information programs are to avoid the cynical judgment of Mr. Dooley, the value of information to the consumer must be our guide.

References

- Charles of the Ritz Distributors Corp. v. Federal Trade Commission, 143 F. 2d. 676 (1944).
- DeKornfeld, Thomas J., Louis Lasagna & Todd M. Frazier. A Comparative Study of Five Proprietary Analgesic Compounds. Journal of the American Medical Association, 1962, 182, 75-78.
- Economic Research Service, Statistical Reporting Service, Consumers and Marketing Service. 1970 Supplement to Statistical Bulletin No. 333, USDA, 1971.
- Friedman, Monroe Peter. Consumer Confusion in the Selection of Supermarket Products. Journal of Applied Psychology, 1966, 50; 529-534.
- Houston, Michael J. The Effect of Unit-pricing on Choices of Brand and Size in Economic Shopping. Journal of Marketing. 1972, 36, 51-54.
- Hutchinson, T. Q. Consumers' Knowledge and Use of Governmental Grades for Selected Food Items. USDA, Economic Research Service, Marketing Research Report No. 876, 1970.
- McCullough, T. David & Daniel I. Padberg. Unit Pricing in Supermarkets: Alternatives, Costs and Consumer Reactions. Search, Agriculture, I (January 1971) 1-25.
- Monroe, Kent B. The Information Content of Prices: A Preliminary Model for Estimating Buyer Response. Management Science, 1971, 17, B519-B532.
- Monroe, Kent B. & Peter J. LaPlaca. What Are The Benefits of Unit Pricing? Journal of Marketing, 1972, 36, 16-22.
- U.S. Department of Agriculture. Food For Us All: The Yearbook of Agriculture, 1969. USDA, 1969.

THE ECONOMICS OF "CONSUMERISM": CAN
COLLECTIVE BARGAINING WORK FOR THE CONSUMERS?

Simone Clemhout¹
Cornell University

The Welfare of the Households

Despite the rise of conventionally defined standard of living indicators (Nordhaus, 1971), the welfare of contemporary households is being jeopardized by a whole battery of factors. In response to such threats, the emergence of "consumerism" is trying to safeguard the individual households in their roles of consumers.

In the best of all possible worlds, classical economists believed that markets operate smoothly, allocating scarce resources to produce goods demanded by the households and the households in turn supply their labor services to the producers. Everything works well as if an invisible hand were keeping the merry-go-around going, with both the standards of efficiency and equity attained.

Soon it transpired that real-life producers exploit their monopsonistic powers in the labor market for their own profit. Both equity and efficiency standards are threatened. Worse still, low wage rates lead to inadequate effective demand and depressions imperil the very foundations of the free-enterprise economy. The rise of labor unions and the advent of management-labor collective bargaining appear to redress somewhat the power imbalance in favor of household as against firms. However, even then statistical evidences suggest that in many industries the U.S. work force is not earning product wage rates commensurate with the marginal products of their labor (Hildebrand & Liu, 1965).

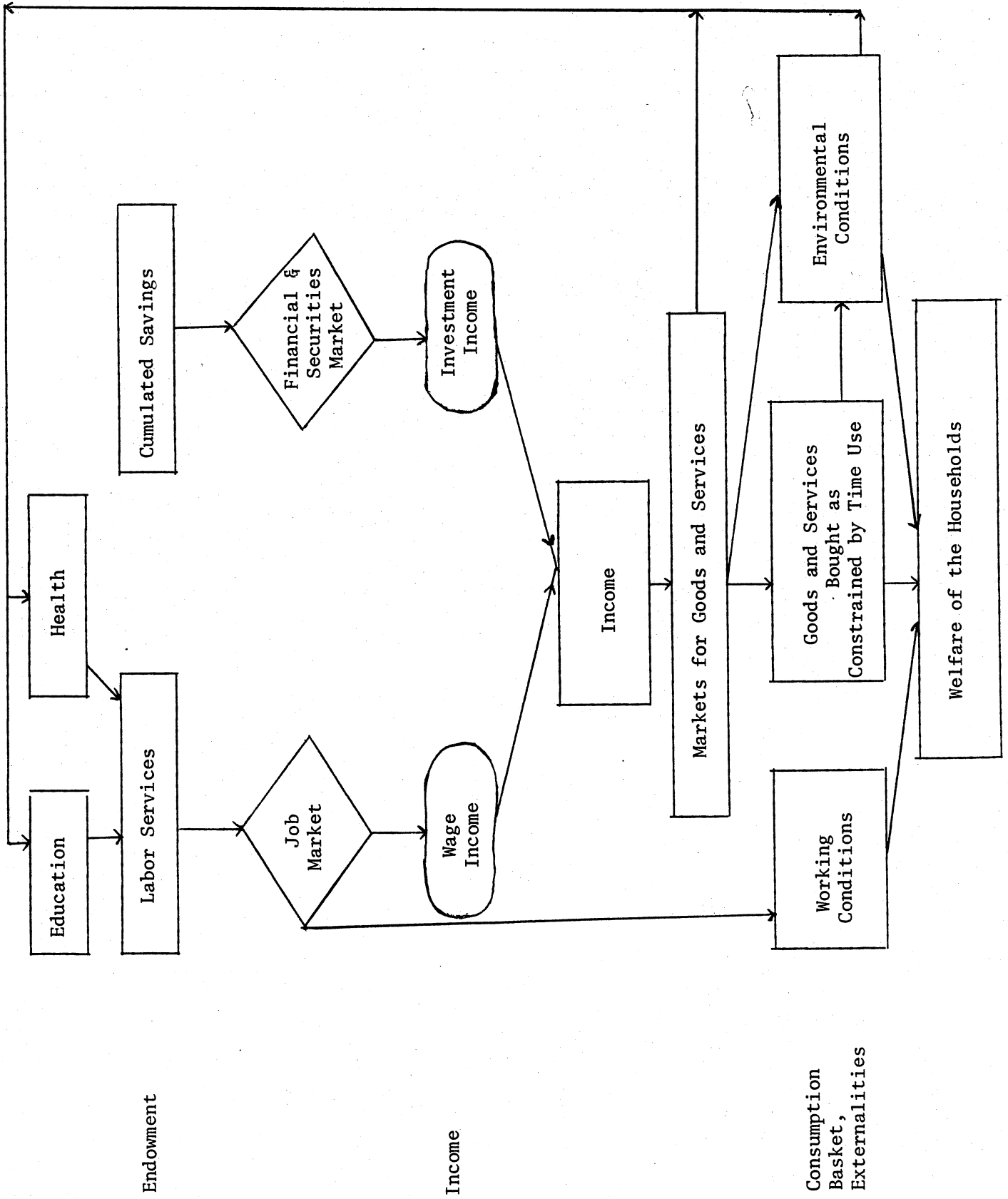
Of late, various signals indicate that the households are still not doing well in their dealings with firms, many of these seem to exploit their market-power in the form of monopolistic power in the product markets. Again, both efficiency and equity are threatened to the detriment of households (in the capacity of consumers). Of course, institutions exist (e.g. Anti-trust laws) to curb the blatant application of monopolistic power. Nonetheless, the situation is far from satisfactory as witnessed by the rising politico-social tides of "consumerism" (Aaker & Day, 1971; Clemhout, 1971; Furuhashi & McCarthy, 1971; Gardner Jones, 1970; Mather, 1971; and Nadel, 1971).

In analogy to the management-labor relationship where collective bargaining worked reasonably well, one may well wonder whether the same collective bargaining approach can be fruitfully transplanted into the producer-consumer framework. What are the feasibility criteria, the drawbacks and the strengths of the collective bargaining approach in the present context, both in its own right and in comparison with the more traditional means of protection for the consumer? To understand the basic issues and to evaluate the alternative policy and institutional measures, we must first analyze the causal determinants of the welfare of the households. The socio-economic system surrounding the household is sketched in Chart 1 below:

As indicated at the bottom of Chart 1, the welfare of the household is affected by three elements:

1. the consumption basket
2. the working conditions
3. the environmental conditions

Flow Chart 1. Factors Determining Households' Welfare



However, to trace the causal chains leading to these three elements, we shall start with the endowments of the households. These include labor services they possess and the cumulative savings they own. By entering the labor market and the financial/security markets, the households receive wage income and investment income. The combination of these two, becomes the household's income which is exchanged for the consumption basket in the market for goods and services. The labor market functions so as to regulate not only the volume but also the conditions of work. The consumption patterns of the household and the factory-and-product designs of the firms determine the environmental conditions. The expenditures of the households are spent not only on the consumption basket but also on health and education services. The environmental conditions not only affect the welfare of the households but also decide--in conjunction with health and education expenditures--the levels of health and education which in turn determine the quality of labor services which a given household can provide. It is now seen that the welfare of households depends upon the interaction of a system of socio-economic factors. It is against this background that we shall appraise the existing methods designed to resolve consumer problems and to investigate the potential of collective bargaining to resolve consumers-producers conflicts.

To protect the welfare of households, there exists already a broad spectrum of measures:

A. Indirect Measures

- (1) To reduce the economic power of the firms--Anti-trust regulations, etc.
- (2) To enhance the economic power of the households--Labor unions, Consumers cooperatives in purchasing.

B. Direct Measures

- (1) General provisions
 - (a) Industry specific: the public regulation of utility industries
 - (b) Problem specific: the pure food-and-drug laws, the anti-pollution ordinances, etc.
- (2) Special provisions--Damage suits by "victimized" households.

For reasons which will be made explicit below, the ensemble of the above institutions and provisions do not constitute an adequate, satisfactory shield to protect the welfare of the households.

Evaluation of Existing Institutional-Policy Measures for the Protection of Consumers' Welfare

In evaluating the alternative institutional/policy measures, we shall first propose three criteria:

- (1) Efficiency versus Equity. The ultimate aim of public policy is presumably an equitable sharing of the economic benefits. While social equity should never be sacrificed on the altar of productive efficiency, likewise any desirable social system should not overlook the efficiency aspect in its pursuit for a more equitable distribution of economic welfare.
- (2) Effectiveness. It is self-evident that any proposed safeguard for individual consumers must be highly effective in fulfilling its avowed aim.
- (3) Viability. Any proposed system must be self-perpetuating if it relies upon the voluntary efforts of individual households.

Table 1 summarizes the comparative strengths of the various safeguarding methods for the consumer.

Table 1

<u>Existing Systems</u>	<u>Desiderata</u>		
	<u>Efficiency</u>	<u>Effectiveness</u>	<u>Viability</u>
1. Reduction of Sellers' Power			
1.1 Present Level		Questionable	
1.2 Intensified Level	Questionable		
2. Public control/public regulation of industries	Questionable	Questionable	Questionable
3. Public control of commodity quality and price	Questionable	Questionable	
4. Judicial supervision		Questionable	Questionable
5. Consumer cooperatives in production			Questionable

At most, anti-trust policies in the past decade, have prevented the take-over of whole industries by monopolists. Oligopolistic firms (e.g., automobile, steel, petroleum industries) do not behave in a manner approximating the ideal of perfect competition. The volume of transactions of such firm is larger than that of consumers by several orders of magnitude. The uneven bargaining strength between buyers and sellers is glaring. On the other hand it is not efficient to reduce sufficiently the seller's market power by breaking-up firms enjoying economies of scale.

The other evident alternative is to substitute profit-motivated privately owned firms with public owned firms or public regulated firms as is done in a number of Western European countries. Not only such solutions are politically alien to the American social will, but judging from the performance of the limited number of U.S. industries which operate upon such basis, e.g., postal service, intra-urban public transportation, etc., there is serious question as to how low the productive efficiency will fall, once the private initiative and profit motive are removed from the system.

The time honored cooperative movement covers two areas of interest:

(1) in the distribution area, it exhibits limited success since we note that of the 3,647 million pounds of the British consumers' expenditure in the year of 1970 only 15 percent are purchased through cooperatives (Nash, 1950).

(2) in the production area, few cooperative enterprises in this country over delve into such activities.

Presumably, the massive capital requirement and complicated technical know-how preclude consumer cooperatives to produce many commodities, e.g., automobiles, T.V. sets, solely for the consumption of their members. Perhaps their comparative advantage does not reside in the production of consumer durables, but products for current consumption like canned and frozen foods. We may therefore conclude that the cooperative movement need not provide an adequate answer.

Likewise, administrative and judicial measures are never likely to safeguard fully the interests of consumers. Who can decide that the price of a certain commodity should fall by how much in view of the latest technical innovation? How can we regulate the desirable durability of a product which arises out of the current fashion change? Major producers' abuses such as poor automobile designs which can be dangerous to drivers, can and should be rectified by legislation and ordinance. By and large, the scope of such measures must necessarily be a limited one. Meticulous regulation is not only impractical but also threatens the efficient operations of the production sector.

The Collective Bargaining Approach

Since the existing measures to protect the consumer have only partial impact and leave unsolved a range of issues we will consider how the collective bargaining approach can be useful in this area.

The prima facie desirability of the collective bargaining approach is supported by the following facts:

(1) The importance of the problems

(a) From the static point of view, on many consumption items, a sizeable portion of the consumer's dollar is spent on distributive services. For instance, of each dollar paid on household appliances, 30 to 40¢ go to the services of wholesalers, retailers, etc. Granted the middlemen perform their share of service, it is still questionable whether the prices for such services are simply too high.

(b) From the dynamic point of view, during the present anti-inflationary campaigns, often times the government cites the stabilization or "rolling back" of the whole-sale index as a sign of improvement, yet there has rarely been any fall of retail prices. Thus, it is the middlemen, rather than the consumer, who benefit from the anti-inflation measures. This can hardly be regarded as an acceptable state of affairs.

(2) The comparative advantage of the approach. The price determination mechanism in a free-enterprise society is as intricate as delicate clockwork. While there exists well-known and significant flaws of this mechanism, most of the present day policy measures are simply too blunt as instruments to rectify the aforementioned flaws. Since the malfunctioning of the system is the consequence of the massive market power concentrated in the hands of the sellers, it is the natural counter-measure for the households to consolidate their market power.

In the commodity markets, few commodities have no close substitutes. On the other hand, modern production technology is such that firms can hardly reorient their product lines without sizeable costs of adjustment. In principle, once "coalitions" can be formed among the buyers, the bargaining advantage of the producers will be threatened. Of course, the crux of the matter is the formation of bargaining coalitions among the buyers.

One may argue that the management-labor collective bargaining has a history of over a hundred years. Since the concept of producer-consumer bargaining is almost symmetric to the former, the failure of the latter to emerge is conclusive evidence that there exist unsurpassable difficulties for such an institution to materialize. While a scientific student will not rule out the validity of the

conjecture, a priori, the logic of such reasoning is far from watertight. Most institutions are viable only under specific circumstances. Furthermore, there always exists a time-lag between a ripen socio-economic environment and emergence of a specific institution. We cannot use past history alone to preclude the possibility or desirability/feasibility of the producer-consumer bargaining in the near future under any form or for any sub-sector of our economy. It is for this reason that we engage in this economic analysis of the subject.

While the concept of coalition enjoys a predominant position in the theory of cooperative games, the usual game-theoretic discussion is on a more general and abstract plane. Hence, some tailor-made analysis is needed for our purpose. In principle the bargaining coalition can win concessions by two different means.

1. By playing one oligopolist against another
2. By threatening a monopolist with boycott

While casual observers may feel the first case is what one would expect, the economic interpretation of the case is by no means trivial. Suppose firm A is ready to make a given price-reduction to lure n buyers in a coalition to stop buying from firm B, it should be equally willing to make the same offer to n buyers who are not in one coalition. So all that a coalition can achieve is to help firm A in assessing the demand elasticity of a given commodity. Although this case may be of some real-life relevance, we shall not elaborate on this aspect of the problem.

Our analysis of the boycott is conducted through the following simplified model.

We shall derive five results:

- (1) bargaining may improve both equity and efficiency
- (2) bargaining can improve equity and efficiency, but not to the extent of either abolishing the monopolistic profit or the deadweight loss
- (3) massive defection of the members of the bargaining coalition may reduce the collective bargaining approach to ineffectiveness
- (4) contrary to the intuition of many, the defection of some of the members from a coalition need not induce a stampede for abandoning the coalition
- (5) the potential gain from the bargaining approach may vary from situation to situation

Consider the case where we have one seller and n identical price taking buyers; each individual's demand curve can be depicted as dd' on Figure 1. For n given price this graph shows three areas of interest: the consumer surplus, the monopolistic profit and the dead weight loss. The bargaining feasibility frontier relevant for our analysis is determined by tracing out a curve representing the trade off between consumer surplus (the payoff for the buyer, π_b) and monopolistic profit (the payoff for the seller, π_s) as the sales price varies. This is shown as FF' in Figure 2.

Elementary price theory indicates that if the sales price is anywhere above the marginal cost a dead weight loss occurs. In fact, it may be shown that by making side-payments from the buyers who purchase at marginal cost to the seller

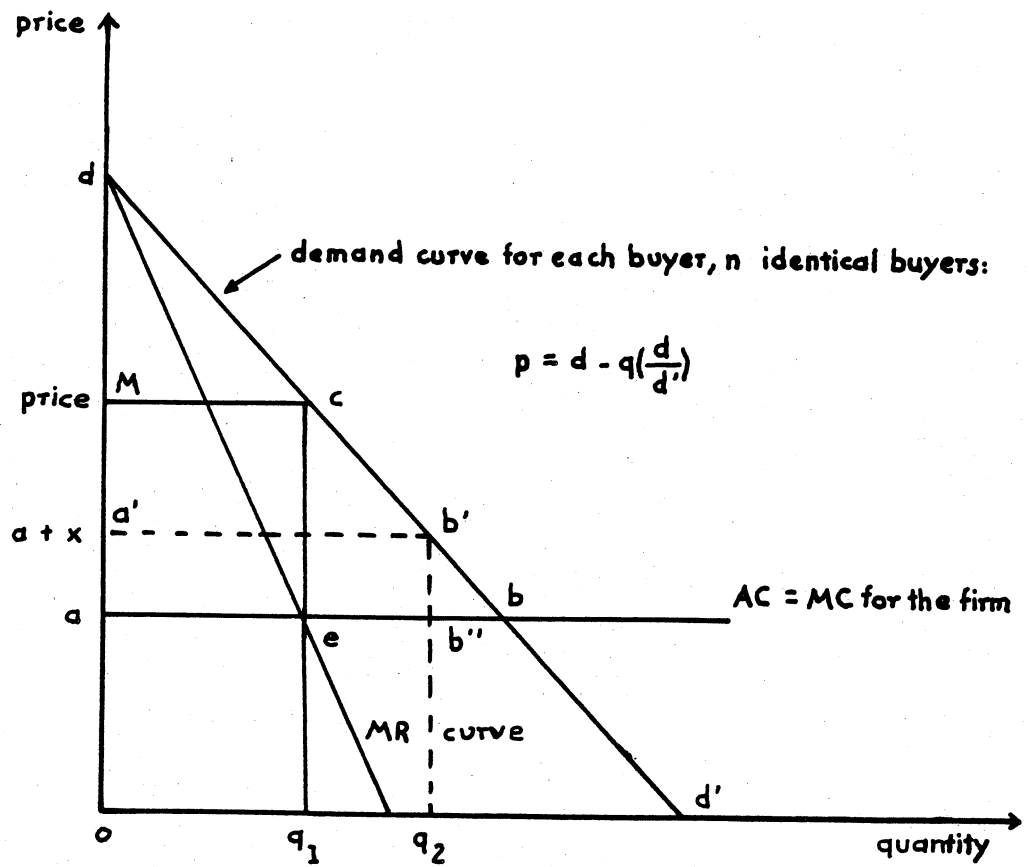


Figure 1

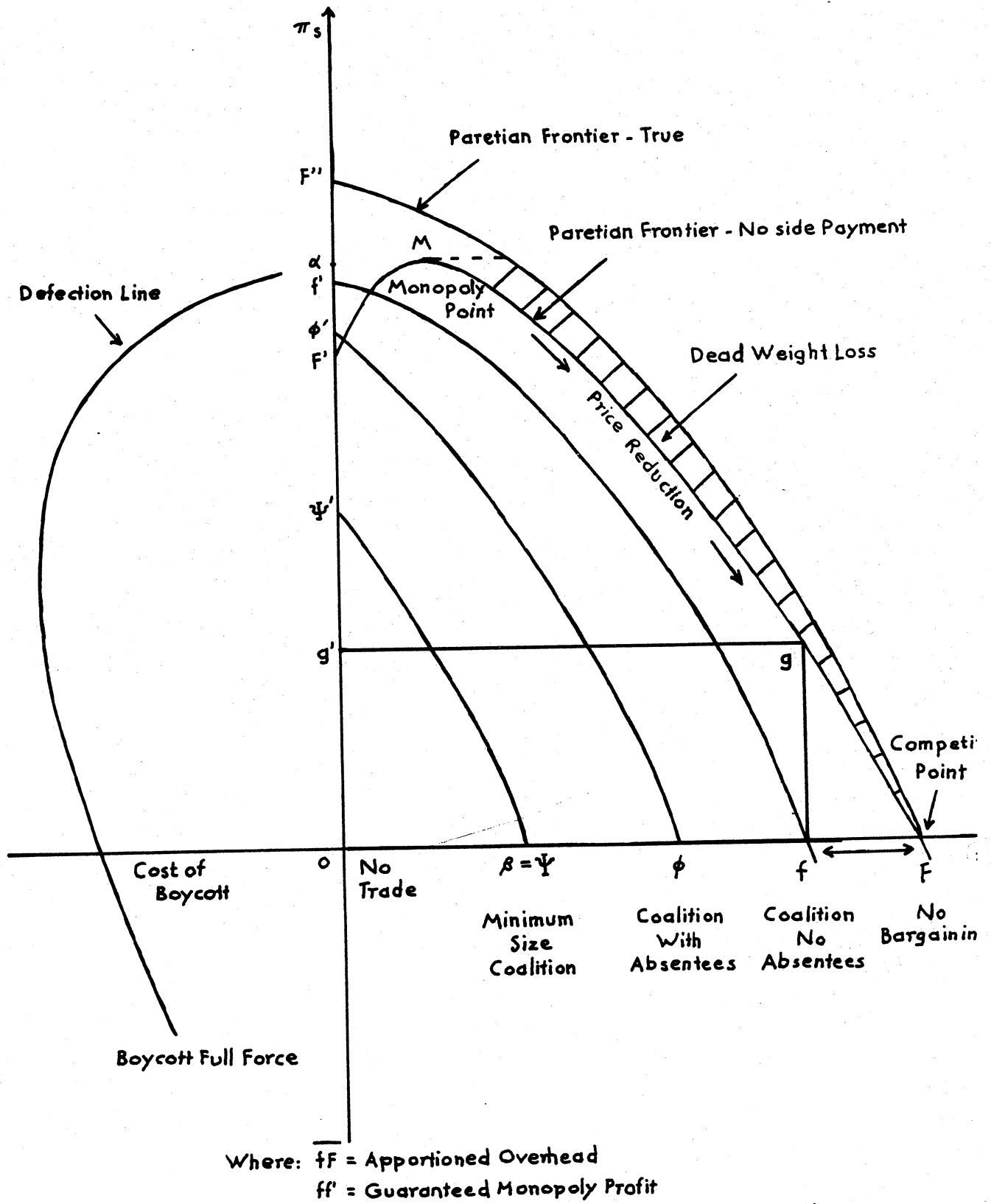


Figure 2

it is possible to obtain a higher frontier FF' denoted as Paratian frontier-true. The horizontal distance between FF' and FF'' represents the dead weight loss. Figure 1 depicts a special example where the demand curve of the representative buyer is linear and the marginal cost, equals the average cost, for the firm is a constant $a > 0$.

The monopolist would maximize his profit by setting unit price = M which induces each buyer to purchase q_1 units. This gives rise to a monopolistic profit of $ceaM$ (in Figure 1) which equals α dollars (in Figure 2) and a consumer surplus of cMd (in Figure 1) which equals β utils (in Figure 2).

Collective bargaining does not involve side-payments but aims at improving the terms of trade for the buyers. However, any successful bargaining implies the reduction of the sales price, accompanied by both the reduction of the monopoly profit and the reduction of the deadweight loss. Should price be reduced from the monopolistic price M to a' , the monopoly profit will fall from $Maec$ to $a'ab''b'$ and the deadweight loss will fall from ceb to $b'b''b'$. Hence both efficiency and (presumably) equity improve with successful bargaining. This justifies result a. We can now consider the limitation of the collective bargaining approach. The setting up of a bargaining unit implies an overhead cost w in terms of utils. If there are $n' \leq n$ buyers participating in the bargaining coalition this would impose a cost of $\frac{w}{n'}$ utils.

The incurring of such cost shifts horizontally the FF' locus to the left since no bargaining coalition wishes to force the seller out of business. At most the maximum payoff for each coalition buyer can be $(F - \frac{w}{n})$ i.e. if every buyer joins the coalition so that the apportioned overhead cost per buyer is reduced to the minimum. This gives rise to a new frontier ff' with the maximum attainable payoff for each buyer = f utils

$$\text{where } \bar{fF} = \frac{w}{n}$$

If there are some buyers absenting themselves from the coalition, the heavier overhead cost per person will shift the feasible payoff frontier further to the left to a position like $\phi\phi'$ and $\psi\psi'$. It now becomes clear that no coalition is worthwhile to pursue if the potential increase of the consumer's surplus for each person is not more than $\frac{w}{n}$.

If the seller wishes to headoff the formation of any bargaining coalition on the side of the buyers, all he has to do is to cut the sales price to $a + x$ (in Figure 1) for some x . [2]

Even if all the n buyers form a coalition, none of them can ever hope to obtain a consumer surplus beyond what they can already achieve under the current unit price $a + x$.

In Figure 2 by constructing a vertical line from f to g on the FF' frontier and then constructing horizontal line gg' we can immediately deduce the minimum attainable monopoly profit og' and the minimum irreducible dead weight loss $bb'b''$. This substantiates result b above. As a corollary to the above analysis, the smaller is the coalition, the higher will be the apportioned overhead $\frac{w}{n'}$ and the more to the left will the feasibility frontier recede. No coalition is worthwhile to pursue if the size of the coalition is so small that the feasible payoff frontier moves to the left of β , namely the consumer surplus which one can obtain anyway under monopoly.

Consequently, it is desirable that the benefit or concession obtained by the bargaining coalition from the sellers be restricted only to the members of the bargaining coalition - no free riders can be allowed to benefit otherwise there may be a universal urge for buyers to absent themselves from the coalition to the point where the coalition shrinks below the minimum survival size, i.e., the feasible payoff frontier retreats to the left of point β .

We now turn to a phenomenon akin to the free riders who dodge the apportioned overhead cost: the defectors who withdraw from a boycott campaign. The former cuts the financial ground from under a bargaining coalition; the latter cuts the real-politic ground from under a coalition in "a state of economic war".

Now in order to force the monopolist seller to make any concessions at all in decreasing its profit below α effective and creditable threats must be presented to the monopolist. In the present context which is collective bargaining without external intervention the only available means is boycott.

We first observe that point o in Figure 2 represents the payoff for both seller and buyer when there is no trade.

In a boycott situation the monopolist loses some or all of its sales while the boycotters must bear not only the loss of consumer surplus but also the necessary expenditure of resource to launch and sustain the boycott campaign. Assuming that n is large a one person boycott would only decrease the monopolistic profit by an iota from the maximum monopolistic profit α . On the other hand with the decrease of the number of boycotting buyers the per person boycotting cost due to organization costs also shrinks towards zero. The payoff locus under boycott is traced out by varying the number of boycotters, it is the curve shown in Figure 2 as defection line. It can be expected that the monopolist seller will make little concessions toward an ineffective boycotting group.

It is also obvious that without the hope of simultaneous action taken by other buyers few individual buyers will launch any boycott at all.

It is the coordinated, organized movement which will sustain the individual urge to boycott. It is also the likelihood of mobilizing a sizeable portion of its membership in a boycotting showdown which can impart any modicum of creditability to the threat of a bargaining unit. All these remarks support our result c .

Since any member in the bargaining coalition may defect in a boycott, one natural question is can the collective bargaining approach ever work for the consumer? In comparison with the situation of labor disputes, it takes picket lines, the exclusive practice of union shops, etc. to knit the union members into a coordinated body in strikes. Consumers have little likelihood to develop similar institutions to back up any boycott. But without boycotts, how much bargaining power to bargain with will the bargaining unit retain? To evaluate the strength of such views, we shall consider the problem through the concept of Nash-equilibrium for non-cooperative games.¹¹ While the coalition problem deals with the cooperative concept, the more fundamental issue of to-join or not-to-join can best be studied along the lines of non-cooperative games. (As Nash noted, non-cooperative games lie at the foundation of cooperative games.) (Nash, 1953).

Assuming there exists no external enforcing means against defection, then each coalition member must weigh the consequences of defection or non-defection. In general, the defecting member can dodge the burden of participating in the

boycott, i.e.: the abstention from trade and the direct boycott cost (organizing, etc.). On the other hand, the act of defection may affect the outcome of the boycott and even the defecting members still want to enjoy the fruits of a successful boycott - a favorable subsequent settlement. Therefore, those members believing that their own participation of the boycott is non-consequential may defect. Others who are convinced of their potentially pivotal role in the final outcome will persevere in the boycott movement. While the possibility of defection weakens the stand of a boycott, it does not necessarily doom such movements to failure, defections notwithstanding. The following example illustrates the situation.

Let n' be the number of buyer in the bargaining coalition calling a boycott.

$n'' < n'$ be the number of non-defecting members required to make the boycott a success, although all n' persons will benefit from the boycott.

π_1, π_2, π_3 and π_4 are the expected payoffs with $\pi_1 < \pi_2 < \pi_3 < \pi_4$ as shown in the following matrix:

Payoff	<u>Boycott successful</u>	<u>Boycott unsuccessful</u>
Defecting	π_4	π_2
Nondefecting	π_3	π_1

$(m, n' - m)$, with $0 \leq m \leq n'$, stands for the defection/non-defection split within the coalition.

Considering now the defection/non-defection as the only two strategies open to all coalition members, it can be shown that the Nash-equilibria are of the following two types: $(n'', n' - n'')$ and $(0, n')$.

The moral is: while it may be difficult to initiate a non-defecting sub-coalition to effect a successful boycott, once a "critical mass" of boycotters have committed themselves, they will stay on to win the issue, even though there are free riding defectors who reap the benefit but dodge the burden of a boycott. In short, there is no intrinsic reason why the failure to punish free riding defectors will prevent a successful boycott from being launched. The above analysis validates our conclusion d.

A three-person special case of the above example is depicted on Figure 3 where the integers 0 and 1 stand respectively for defection and non-defection and it takes two members to organize a successful boycott. The eight vertices are assigned with ordered triplets of payoffs to the three players of the game and any arrow sign between two vertices shows how one unstable "split" tends to transform into another "split" be it stable or unstable.

Conclusion e is rather self-evident. The outcome and benefit of the bargaining process varies from case to case. Our Appendix sketches how these variations affect the result of bargaining under the Nash bargaining model (Nash, 1950).

Some Speculative Comments

What circumstances favor the collective bargaining approach most? From our theoretic analysis, it seems that the following criteria are important:

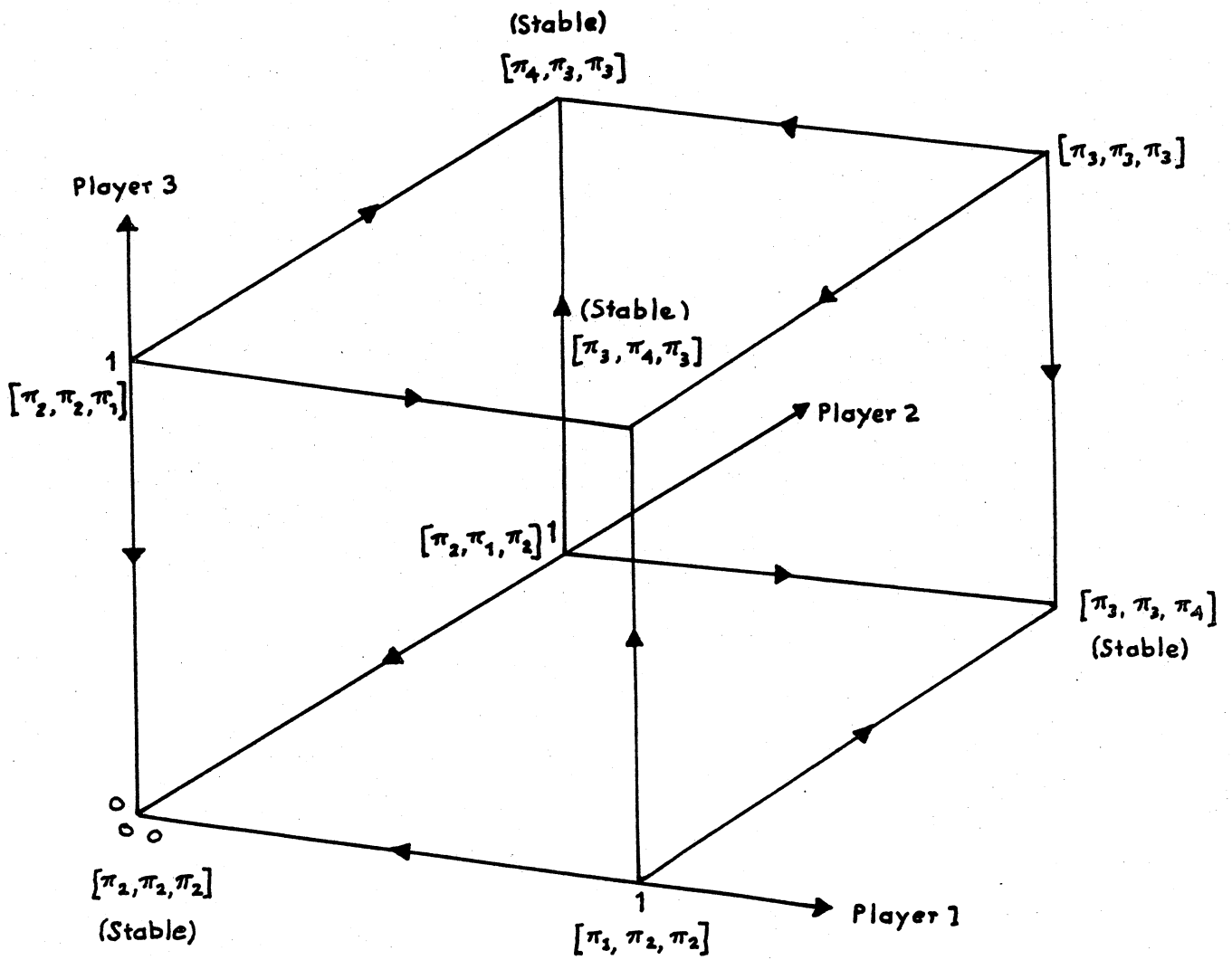


Figure 3

(1) The strength of group-identification on the side of consumers. This factor predicts how extensively the bargaining coalition can control the buyers' actions whether to join the coalition or to launch a boycott.

(2) The importance of a commodity in the consumer's budget. More likely the consumers will be highly motivated for action regarding important commodities.

(3) Substitutibility of the boycotted product. The higher is the elasticity, the easier it is to sustain the no-trade situation under a boycott.

(4) The shiftability of supply. This reflects the power to resist pressure on the side of the sellers. If the shiftability is low, then the seller is not likely to face a bargaining coalition with much resolve.

Social goals can be achieved through collective bargaining which go beyond the hope of traditional Pigovian welfare economics which regards the lowering of monopolistic prices as always desirable since this increases the consumers' surplus. Today equitable income and wealth distribution is regarded as too uneven and, if, due to practical difficulties (e.g. in formulating and passing the appropriate legislation, in administering the required programs, etc.), one cannot have effective and appropriate means to rectify the situation, there is every justification for the disadvantaged, the needy and the poor to improve their economic status through the coordinated utilization of their inherent bargaining power. Such groups of individuals may include the low-income pensioners, the ghetto residents, those populating the economically depressed regions and perhaps part of the student population who cannot quite utilize the financial institutions to tap their prospective earning power through long-term loans.

What are the institutional set-ups around which the collective bargaining units may be found? Taking a leaf out of the history of industrial organizations, it seems it is much easier for the established institutions to take over new functions than for new institutions to be established solely for the purpose of carrying out these new functions. The ready-made constituency for the bargaining coalition, the pre-existing staff and organization which may well have excess capacity to be explored and the high motivation of the membership backed by years of history. If the above observations are correct, then presumably existing labor unions, urban ghetto organizations, students' associations, consumer cooperatives can become the springboard of collective bargaining campaigns (Jackson & Taylor, 1969).

Footnotes

1. The author is Associate Professor of Consumer Economics at Cornell University.
2. The calculation of x is as follows:

In Figure 1, the most each consumer can hope for is to force the seller to sell at cost, a . The maximum attainable increment of his consumer surplus is the area of the trapezoid $aa'b'b$, i.e.,

$$\text{Area of the trapezoid} = \frac{x}{2} [(d-a) + (d-a-x)] \frac{d'}{d}$$

On the other hand, the minimum apportioned cost is $\frac{w}{n}$. If $\frac{x}{2} [2(d-a)-x] \left(\frac{d'}{d}\right) = \frac{w}{n}$

then, there cannot be any coalition. Solving for x from the above equation:

$$x = (d-a) \pm \sqrt{(d-a)^2 - 2\left(\frac{d'}{d}\right)^2 \frac{w^2}{n^2}}$$

Since $a + x \leq d$ (otherwise, quantity of output = $\frac{d'}{d} (d-a-x) < 0$):

$$x = (d-a) - \sqrt{(d-a)^2 - 2\left(\frac{d'}{d}\right)^2 \frac{w^2}{n^2}}$$

Appendix

Nash Equilibrium for the n-person cooperative game

Feasibility Frontier, Nash Solution and Sensitivity Analysis. From Figure 1.

$$p = d \left(1 - \frac{q}{d}\right) \text{ or } \frac{p}{d} = 1 - \frac{q}{d}, \text{ or } \frac{q}{d} = 1 - \frac{p}{d}$$

$$\begin{aligned} \text{Consumer's surplus (per buyer): } y &= 1/2 (d - p)q \\ &= (dd'/2) (1 - p/d)^2 \end{aligned}$$

$$\begin{aligned} \text{Producer's profit (per buyer): } z &= (p - a)q \\ &= (p - a)d'(1 - p/d) \\ &= dd'(p/d - a/d)(1 - p/d) \end{aligned}$$

$$\begin{aligned} \text{Feasible payoff frontier: } z &= dd'(1 - a/d - \sqrt{2y/dd'}) \sqrt{2y/dd'} \\ &= (\sqrt{2dd'} - \sqrt{2a(d'/d)} \sqrt{y} - 2y) \\ &= v \sqrt{y} - 2y, \text{ say} \end{aligned}$$

which is an elipsoidal arc shown as in Figure 4.

There is no consensus about where the bargaining solution may end up. The Nash solution for the bargaining game is regarded as a norm for arbitration by Luce and Raiffa⁸ and as a first approximation to a realistic prediction by Harsanyi and Selten. Without considering the intricacies developed by the elaborate models of Harsanyi and Selten (Harsanyi & Selten, 1969), we shall concentrate on the Nash solution. Moreover, simplifications will be made by assuming:

(1) There is neither direct cost in organizing bargaining by the consumers nor any shutting-down cost by the producers.

(2) All buyers join the boycott.

Tangency between the feasible payoff frontier and the highest equilateral hyperbola $yz = K$ implies the Nash solution (y^*, z^*) satisfying the conditions:

$$K = vy^{*3/2} - 2y^{*2}$$

$$\frac{-K}{y^{*2}} = \frac{v}{2\sqrt{y^*}} - 2 \text{ or } -K = \frac{v}{2} y^{*3/2} - 2y^{*2}$$

Adding the two equations to eliminate K,

$$\frac{3v}{2} y^{*3/2} = 4y^{*2} \text{ or,}$$

$$y^* = \frac{9}{64} v^2$$

$$z^* = v^2(3/8 - 9/32) = \frac{3}{32} v^2$$

Eliminating v^2 , one obtains that:

$2y^* = 3z^*$ for all v , i.e., for all a, d, d' as shown in the above figure as the settlement path.

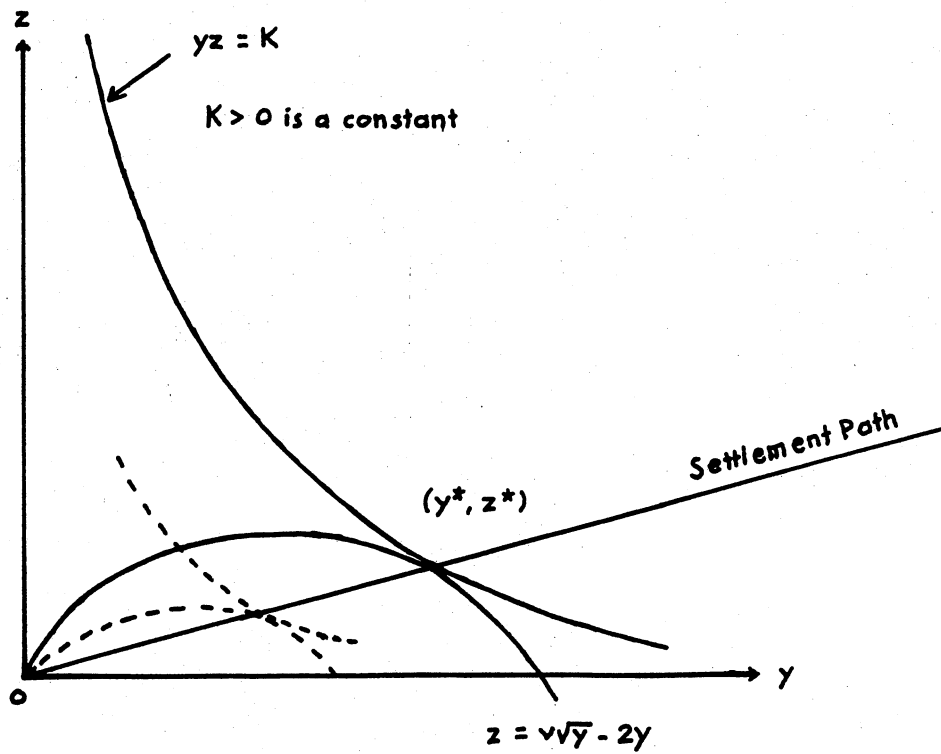


Figure 4

References

- Aaker, D. A. & George S. Day. Consumerism: Search for the Consumer Interest. New York: The Free Press, 1971.
- Clemhout, Simone. Consumer Disaffection: Time to Bargain? Human Ecology Forum, 1971, 2.
- Furuhashi, Y. H. & McCarthy, E. J. Social Issues of Marketing in the American Economy. Columbus, Ohio: Grid, Inc., 1971.
- Gardner Jones, M. Wanted: A New System for Solving Consumer Grievances. Arbitration Journal. 1970, 25, 234-347.
- Harsanyi, J. C. & Reinhard Selten. A Generalized Nash Solution for Two-Person Bargaining Games with Incomplete Information. Working paper No. 285, Center for Research in Management Science, University of California, Berkeley, October 1969.
- Hildebrand, G. H. & T. C. Liu. Manufacturing Production Functions in the U.S. 1957. Ithaca, New York: Cornell University Press, 1965.
- Jackson, Samuel C. & Warren L. Taylor. Center for Dispute Settlement - A New Use for an Old Tool. Legal Aid Briefcase, February 1969, 130-135.
- Luce, D. R. & Howard Raiffa. Games and Decisions: Introduction and Critical Survey. New York: John Wiley and Sons, Inc., 1957.
- Mather, L. L. (editor). Economics of Consumer Protection. Danville, Illinois: The Interstate, Printers and Publishers, Inc., 1971.
- Nadel, Mark V. The Politics of Consumer Protection. New York: Bobbs-Merrill Co., Inc., 1971.
- Nash, J. F. Non-Cooperative Games. Annals of Mathematics. 1951, 54, 286-295.
- Nash, J. F. Two-Person Cooperative Games. Econometrica. 1953, 21, 128-40.
- Nash, J. F. The Bargaining Problem. Econometrica. 18, 1950, 1955-1962.
- Nielsen Reports. Co-ops Retain Share of Grocery Trade. Cooperative Management and Marketing. 1971, 4, 33-35.
- Nordhaus, W. GNP and the Quality of Life. Presented at American Academy for the Advancement of Science, Boston, December 30, 1971.

MARGINAL SALIENCE OF PRICE IN BRAND EVALUATIONS¹

Vithala R. Rao
Cornell University

Introduction

The various brands in any product class can be described as a set of multi-attribute alternatives. These attributes may include both the significant and symbolic aspects (Coombs, Dawes, & Tversky, 1970). Consumer's decision to choose a brand can be thought of as a composite decision rule incorporating the various attributes of the brands. These rules obviously reflect the individual's idiosyncratic utility function.

One critical attribute of the brand that has been intensively studied by economists is brand price. It is only in the past decade or so that researchers of consumer behavior in marketing turned their attention to studying price as a perceptual dimension of evaluations with respect to brand quality and brand worth (defined as some measure of quality per unit price.) Past research (Gabor & Granger, 1966, Gardener, 1971, Jacoby, 1970, Leavitt, 1954, McConnell, 1968 a & b, Peterson, 1970, Rao, 1972, and Tull, Boring & Gonsoir, 1964) has indicated that price is used by consumers as a surrogate for quality in the absence of other brand information and that the importance of price in quality perceptions diminishes when a number of other brand cues are present.

Thus it appears that the salience of price is not independent of other brand information. This paper is a preliminary attempt to determine the nature and magnitude of the trade-off relations between importance of price and the number of non-price informational cues. The research question under study is whether there exists an optimal number of additional brand cues in the presence of which price is least salient. Obviously, the existence of such a number has pragmatic implications for promotional planning and public policy. The paper is organized into three broad sections. (1) formulation of the hypothesis and discussion of the measurement model employed; (2) discussion of the experimental design and analysis procedure; and (3) results and possible implications.

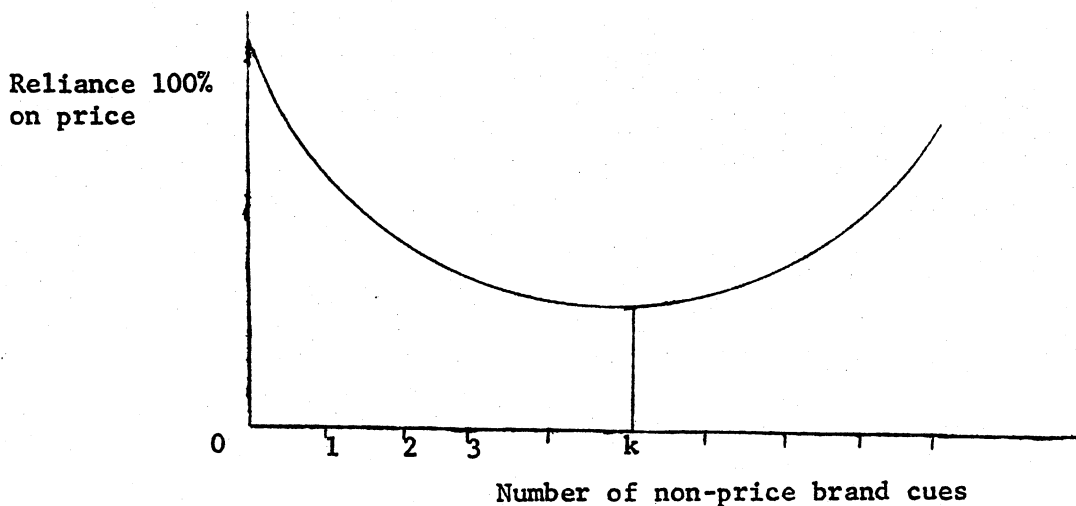
Hypothesis and Measurement

The substantive hypothesis under examination is that consumer's reliance on price in making overall brand evaluations is nonlinearly related to the number of additional non-price informational cues on brands. As previous research indicates, not surprisingly when no other information is present consumer rely solely on price in their brand evaluations. Furthermore, as information on other attributes becomes available, they tend to utilize such information thereby decreasing the importance assigned to price. However, when the additional cues are too many (beyond the individual's cognitive limits with respect to information overload), the individual will tend to revert to relying upon generally accepted index such as price. That is to say, the importance assigned to price will increase

after a certain number of additional non-price cues. This hypothesis may be shown as in Figure 1. The curve depicts the way in which importance on price is related to the number of additional cues. It reaches a minimum at the point k additional cues and rises thereafter.

Figure 1

HYPOTHESIZED RELATIONSHIP BETWEEN
RELIANCE ON PRICE AND NUMBER OF
NON-PRICE BRAND CUES



The model of additive conjoint measurement, developed by mathematical psychologists, (Coombs et. al., 1970, Luce, 1964, Roskam, 1968, Tversky, 1967 a & b) is employed in this paper for measuring the importance of price. This model is based on a paradigm where an ordering of a dependent variable (namely overall evaluation of preference) is obtained under different combinations of two (or more) independent variables (price and non-price attributes). To make the discussion concrete, let us consider the product class of automobiles and a set of $m \cdot n$ hypothetically described auto stimuli on two attributes (A) size and (P) price respectively, of m and n levels and identical in all other respects. Each stimulus has two coordinates (a_i, p_j) , $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$. Assume that the mn stimuli have been ordered by an individual with respect to his overall opinion of worth of the automobile to him. Let $M(a_i, p_j)$ represent the ordinal measure of his evaluation. The additive model of conjoint measurement assumes that there exist functions f , g , and ϕ defined on $A = (a_i, i = 1, 2, \dots, m)$; $P = (p_j, j = 1, 2, \dots, n)$ and AXP respectively, such that

1. $\phi(a_i, p_j) = f(a_i) + g(p_j)$
2. $\phi(a_i, p_j) \geq \phi(a_k, p_l)$ if and only if $M(a_i, p_j) \geq M(a_k, p_l)$

Thus, if the model holds, the cell entries of M-matrix can be rescaled such that their order is preserved and such that every rescaled entry is expressed as the sum of the components of the two attributes, A(size) and P(price). Such a model exists if the axioms of cancellation and solvability are satisfied (Coombs et. al., 1970). Under these conditions, the two attributes A and P can be regarded as independent in the sense that they contribute independently (additively) to produce the joint effect. This model is quite similar to that of the analysis of variance with no significant interactions. The variation in rescaled values of M-matrix can be written as the sum of the variations contained in the component functions (hereafter called partworth functions) of the attributes A and P. The operational measure of the importance of price in overall judgments is then the percent of the total variation in rescaled values of M due to attribute, P:

$$\text{Importance of price} = \frac{\text{Sum of squares attributable to price}}{\text{Total sum of squares of rescaled M-matrix}}$$

Computer programs exist to perform the above analysis on ranked judgments of stimuli derived according to a factorial design. The program developed by Kruskal (Kruskal, 1965), known as MONANOVA which does monotone analysis of variance is employed in this study. This program computes² a measure of goodness to fit, known as stress (S) defined by:

$$S = \frac{\sum_i (z_i - z_i(\beta))^2}{\sum_i (z_i(\beta) - \bar{z}_i(\beta))^2}$$

where z_i denotes monotonically transformed values of the dependent variable (i.e., ranked judgments); $z_i(\beta)$ denotes the parameters of the fitted model and $\bar{z}_i(\beta)$ denotes the mean of $z_i(\beta)$. The lower limit of this measure (representing perfect fit) is zero.

Experimental Design and Analysis

To understand the dynamics of the trade-off between price and non-price attributes, hypothetically stated brands were used as stimuli. Such a decision avoided the confounding, if any, due to the image conjured up by brand names. The stimuli were automobiles. Each stimulus was described as a vector of attributes.

Based on a preliminary investigation, six attributes (including price) of the automobile considered most important in brand choice were selected. In addition to price, these were: (A) size; (B) horsepower; (C) miles per gallon; (D) repair record; and (E) origin of manufacture. Except for the origin of manufacture, which was described at 2 levels, four levels were selected for the attributes. These are shown in Table 1.

TABLE 1
LEVELS SELECTED FOR ATTRIBUTES

Attribute	Number of Levels	Level
A. SIZE	4	FULL SIZE INTERMEDIATE COMPACT MINI
B. HORSEPOWER	4	60 hp 110 hp 190 hp 270 hp
C. MILES PER GALLON	4	10 15 20 25
D. REPAIR RECORD	4	EXCELLENT GOOD AVERAGE POOR
E. ORIGIN OF MANUFACTURE	2	AMERICAN FOREIGN
P. PRICE	4	\$1,795 \$2,495 \$3,395 \$3,995

In addition to price (P), four combinations of attributes, namely A and B; A, B, and C; A, B, C, and D; and A, B, C, D, and E were used in developing stimulus descriptions. For each combination, 16 hypothetical stimuli were generated using the principles of a Graeco-Latin Square design. For the first three sets, the attributes were combined using each of the four levels making sure that the successive designs were orthogonal to each other. For the fourth set, 1/4 factorial was selected employing the two extreme levels for the five attributes other than origin of manufacture for which the two levels (American and Foreign) were used. For each combination, the stimuli were described without and with price. The design thus enabled comparison of the price influence, not only within each combination but across the various combinations. The layout of the stimuli is presented in Table 2. Each stimulus was described on a 4" x 5" card as a profile description of an automobile. Subjects were presented two decks of 16 stimulus cards each in the experiment.

Four different groups of subjects³ (residents of the Ithaca area including some graduate students of Cornell University) participated in this experiment. Respectively 24, 36, 27 and 31 subjects (totalling 118) responded to the 2, 3, 4 and 5 attribute stimuli. The experimental tasks included the following:

- (a) Ranking of the 16 automobile descriptions without price data from 'excellent' to 'poor' buy for the money assuming the cars were identical on all other characteristics including price.
- (b) Ranking of the 16 automobile descriptions with price information including from 'excellent' to 'poor' buy for money assuming that cars were all identical on all other characteristics.

The ranked judgments of worth of the car excluding and including price information were analyzed for each individual using the Kruskal's method of monotone analysis of variance, known as MONANOVA (Kruskal, 1965). The analysis yielded partworth (utility) functions for each attribute included in the set. For example in the 2-attribute group, analysis of ranks yielded partworth functions for the attributes of size, horsepower, using the first set of ranks and the partworth functions for price also, using the second set of ranks. From these functions, the proportion of total variation due to each attribute before and after price was computed. The measure of reliance on price was simply the proportion of variation attributable to price ranked preferences obtained after price information.

Based on these individual analyses, subjects were divided into two groups: consistent and others. The consistent subjects were those satisfying the two criteria: (a) satisfactory fit of the model to their data with stress levels of 0.10 or below; and (b) the monotonicity of their partworth function of price (i.e., high values of partworth for low values of price and decreases in the expected direction). The number of such consistent subjects in the 4 groups were respectively 14 out of 24; 21 out of 36; 21 out of 27 and 20 out of 31. Further summarization of data was done separately for consistent subjects and all subjects, discussion being confined mainly to the former group.

TABLE 2
STIMULUS DESCRIPTIONS FOR THE 4 GROUPS

Stimulus Number	Number of Attributes																	
	2			3				4				5						
	A	B	P	A	B	C	P	A	B	C	D	P	A	B	C	D	E	P
1	1	1	1	1	1	1	1	1	1	1	1	1	4	1	1	1	1	4
2	1	2	2	1	2	2	2	1	2	2	2	2	1	4	1	1	1	1
3	1	3	3	1	3	3	3	1	3	3	3	3	1	1	4	1	1	1
4	1	4	4	1	4	4	4	1	4	4	4	4	1	1	1	4	1	1
5	2	1	2	2	1	3	2	2	1	3	4	2	1	1	1	1	2	4
6	2	2	1	2	2	4	1	2	2	4	3	1	4	4	4	1	1	1
7	2	3	4	2	3	1	4	2	3	1	2	4	4	4	1	4	1	1
8	2	4	3	2	4	2	3	2	4	2	1	3	4	4	1	1	2	4
9	3	1	3	3	1	4	3	3	1	4	2	3	4	1	4	4	1	4
10	3	2	4	3	2	3	4	3	2	3	1	4	4	1	4	1	2	1
11	3	3	1	3	3	2	1	3	3	2	4	1	4	1	1	4	2	4
12	3	4	2	3	4	1	2	3	4	1	3	2	1	4	4	4	1	4
13	4	1	4	4	1	2	4	4	1	2	3	4	1	4	4	1	2	4
14	4	2	3	4	2	1	3	4	2	1	4	3	1	4	1	4	2	4
15	4	3	2	4	3	4	2	4	3	4	1	2	1	1	4	4	2	1
16	4	4	1	4	4	3	1	4	4	3	2	1	4	4	4	4	2	1

Note: For Level codes, see Table 1.

Results

The average partworth functions derived from the preference judgments excluding and including price for the consistent and all subjects are presented in Tables 3-6 for the 4 groups responding to 2, 3, 4, and 5 non-price attributes. The same data are also graphically represented for the consistent subjects in Figures 2-5.

The following descriptive observations based on the spread and direction of the partworth functions of the various attributes for the consistent subjects may be noted:

- (a) In all cases, size is the least salient factor in the total preference. In general, the subjects evince low preference (utility) for full and mini sizes in their judgments excluding price. The pattern after price is monotonically decreasing with the size level.
- (b) Subjects' utility for horsepower is highest in the 2 and 3 attribute group in their judgments excluding price. In the 4 and 5 attribute case, the horsepower continues to be high, but is dominated by repair record. Furthermore, the pattern is about the same in judgments including price. The partworth function is nonlinear and is increasing with horsepower becoming asymptotic after 250 h.p. or so.
- (c) The function for gas mileage is monotonic with mpg. as expected.
- (d) The attribute of repair record dominates others in 4 and 5 attribute cases. Its specific function is monotonically decreasing with the quality of repair record, as should be expected.
- (e) The origin of manufacture does not turn out to be a salient factor. However, subjects express higher utility for foreign made cars than domestic ones.
- (f) Not surprisingly, in all cases the utility expressed for price attribute diminishes with price.

In order to obtain a quantitative measure of importances assigned to each attribute, the percent contributions of each attribute, the percent contributions of each were computed. As noted earlier these are the percent total sum of squares due to each factor. The means and standard errors of these statistics applicable to the consistent subjects are presented in Table 7. Figure 6 shows the percent reliance on (contribution of) price as a function of number of non-price attributes. While no formal statistical tests are possible, the figure does provide empirical support for the hypothesis under study. The percent reliance on price does diminish monotonically up to the 4 non-price attributes and rises at 5.

Fortunately, owing to the nature of the design, similar analysis is possible for two of the non-price attributes of size and horsepower. In the case of size, the percent reliance decreases steadily from 11.5 percent for the 2 attribute group to 5.5 percent for the 5 attribute case. The pattern is more pronounced for the horsepower attribute the range being 42.7 percent to 15.0 percent. Thus, the importance given to these non-price variables drops steadily with the number of other attributes not showing any signs of recovery as noted for the price variable. This finding enhances the face validity of the hypothesis under examination.

TABLE 3
 AVERAGE PARTWORTH FUNCTIONS FOR JUDGEMENTS
 (EXCLUDING AND INCLUDING PRICE) OF CONSISTENT
 AND ALL SUBJECTS IN 2 ATTRIBUTE GROUP

Attribute	Level	Consistent Subjects		All Subjects	
		Average Partworth Contribution		Average Partworth Contribution	
		Excluding Price	Including Price	Excluding Price	Including Price
Size	Full	-0.28	0.48	-0.28	0.44
	Inter- mediate	0.14	0.17	0.13	0.13
	Compact	0.16	-0.17	0.19	-0.11
	Mini	-0.03	-0.49	-0.03	-0.45
Horse- power (hp)	60	-1.58	-1.42	-1.56	-1.51
	110	-0.08	-0.47	-0.06	-0.46
	190	0.72	0.83	0.73	0.77
	270	0.94	1.06	0.89	1.19
Price(\$)	1,795		1.28		1.01
	2,495		0.67		0.52
	3,395		-0.72		-0.50
	3,995		-1.24		-1.04

TABLE 4
 AVERAGE PARTWORTH FUNCTIONS FOR JUDGEMENTS
 (EXCLUDING AND INCLUDING PRICE) OF CONSISTENT
 AND ALL SUBJECTS IN 3 ATTRIBUTE GROUP

Attribute	Level	Consistent Subjects		All Subjects	
		Average Partworth Contribution		Average Partworth Contribution	
		Excluding Price	Including Price	Excluding Price	Including Price
Size	Full	-0.09	0.49	-0.01	0.38
	Inter- mediate	0.09	0.23	0.05	0.24
	Compact	0.06	-0.33	0.08	-0.23
	Mini	-0.06	-0.40	-0.12	-0.39
Horse- power (hp)	60	-1.28	-1.49	-1.17	-1.48
	110	-0.10	-0.33	-0.08	-0.28
	190	0.53	0.57	0.46	0.62
	270	0.85	1.25	0.80	1.15
Mileage (mpg)	10	-1.55	-0.95	-1.69	-1.11
	15	-0.20	0.04	-0.21	0.07
	20	0.46	0.20	0.60	0.24
	25	1.29	0.71	1.29	0.81
Price (\$)	1,795		1.62		1.29
	2,495		0.43		0.41
	3,395		-0.55		-0.53
	3,995		-1.49		-1.18

TABLE 5
 AVERAGE PARTWORTH FUNCTIONS FOR JUDGEMENTS
 (EXCLUDING AND INCLUDING PRICE) OF CONSISTENT
 AND ALL SUBJECTS IN 4 ATTRIBUTE GROUP

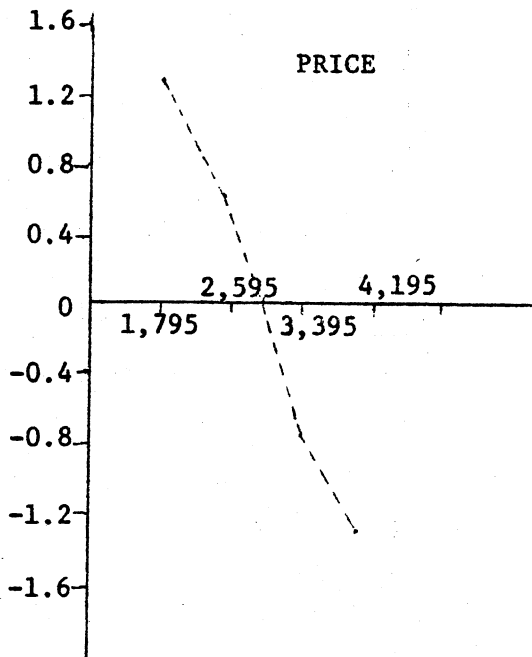
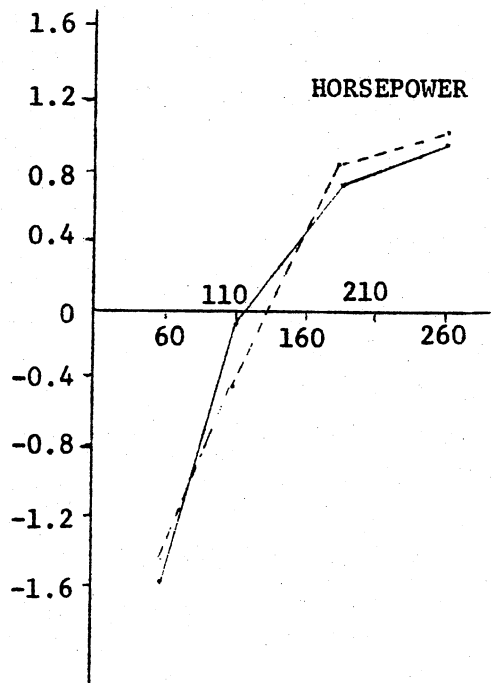
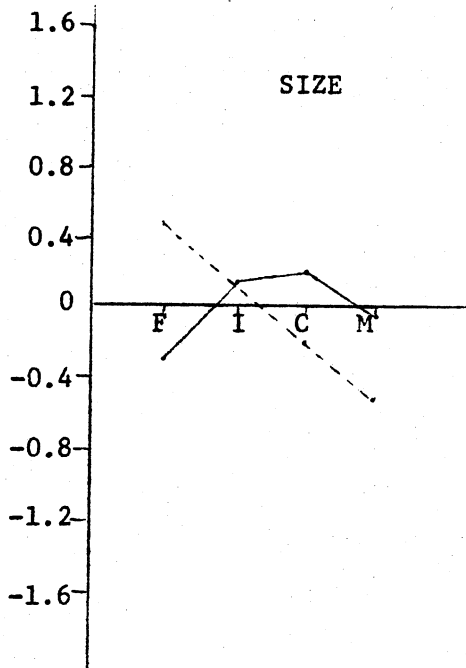
Attribute	Level	Consistent Subjects		All Subjects	
		Average Partworth Contribution		Average Partworth Contribution	
		Excluding Price	Including Price	Excluding Price	Including Price
Size	Full	-0.07	0.46	-0.07	0.29
	Intermediate	0.22	0.17	0.25	0.25
	Compact	-0.02	-0.51	-0.03	-0.51
	Mini	-0.13	-0.13	-0.15	-0.03
Horse-power (hp)	60	-0.77	-1.17	-0.88	-1.32
	110	0.17	0.10	0.25	0.06
	190	0.44	0.40	0.40	0.53
	270	0.16	0.67	0.23	0.72
Mileage (mpg)	10	-1.53	-1.18	-1.45	-1.23
	15	-0.13	-0.08	-0.16	-0.09
	20	0.72	0.48	0.66	0.56
	25	0.93	0.78	0.95	0.75
Repair record	Excellent	1.43	0.99	1.33	0.99
	Good	0.53	0.53	0.61	0.53
	Average	-0.07	0.10	-0.05	0.05
	Poor	-1.90	-1.62	-1.89	-1.56
Price (\$)	1,795		1.51		1.33
	2,495		0.58		0.46
	3,395		-0.47		-0.39
	3,995		-1.62		-1.40

TABLE 6

PARTWORTH FUNCTIONS FOR JUDGMENTS (EXCLUDING AND INCLUDING PRICE) CONSISTENT AND ALL SUBJECTS IN 5 ATTRIBUTE GROUP

Attribute	Level	Consistent Subjects		All Subjects	
		Average Partworth Contribution		Average Partworth Contribution	
		Excluding Price	Including Price	Excluding Price	Including Price
Size	Full	-0.14	0.19	-0.14	0.21
	Mini	0.14	-0.19	0.14	-0.21
Horse-power (hp)	60	-0.44	-0.66	-0.41	-0.68
	270	0.44	0.66	0.41	0.68
Mileage (mpg)	10	-0.70	-0.51	-0.81	-0.52
	25	0.70	0.51	0.81	0.52
Repair record	Excellent	1.71	1.20	1.57	1.19
	Poor	-1.71	-1.20	-1.57	-1.19
Origin of manufacture	American	-0.50	-0.30	-0.60	-0.34
	Foreign	0.50	0.30	0.60	0.34
Price (\$)	1,795		1.40		1.27
	3,995		-1.40		-1.27

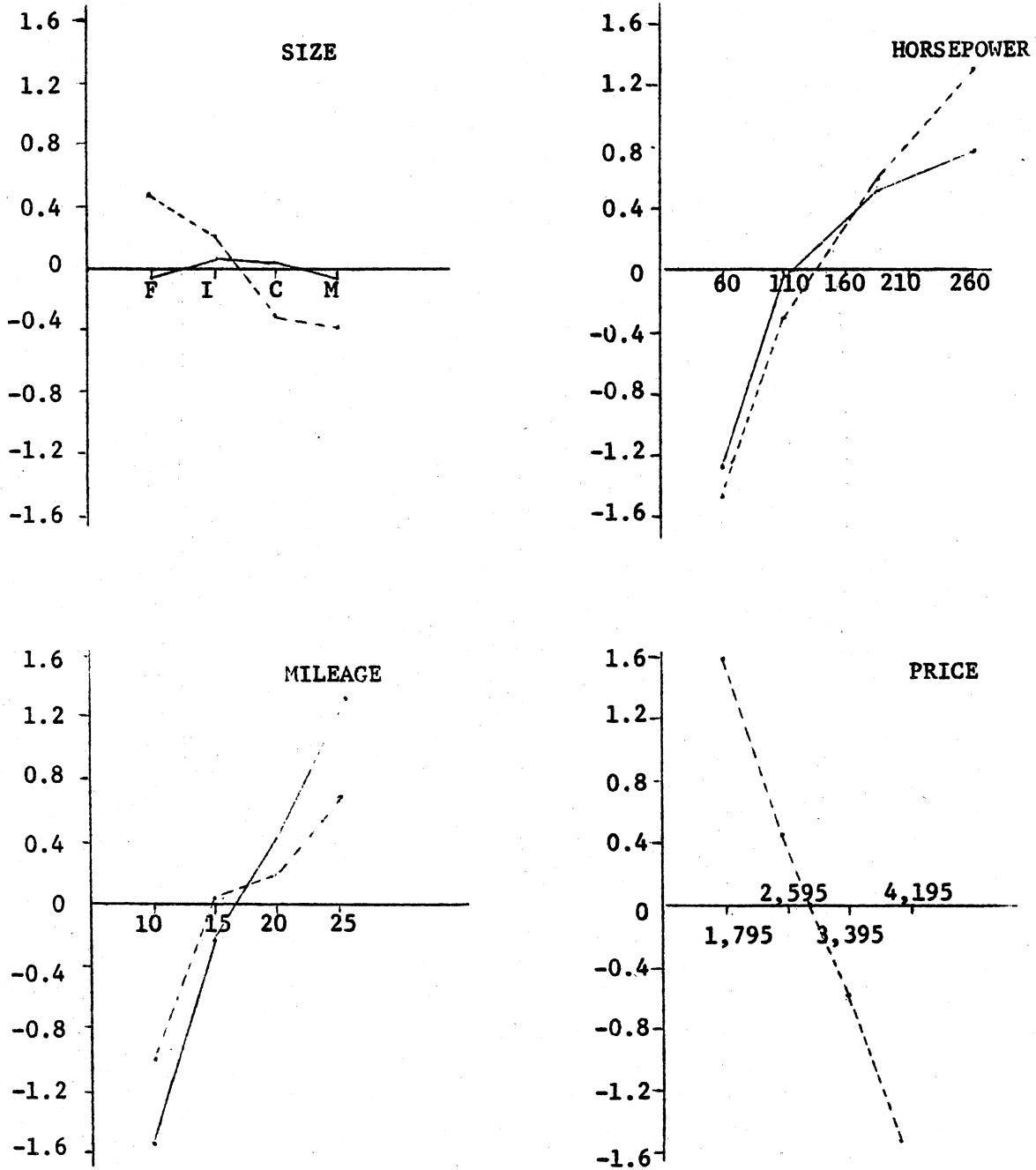
FIGURE 2
 PARTWORTH FUNCTIONS FOR JUDGMENTS OF
 CONSISTENT SUBJECTS IN 2 ATTRIBUTE GROUP



Legend:

- Excluding price
- - - Including price

FIGURE 3
 PARTWORTH FUNCTIONS FOR JUDGMENTS OF CONSISTENT
 SUBJECTS IN 3 ATTRIBUTE GROUP



Legend:

- Excluding price
- - - Including price

FIGURE 4
 PARTWORTH FUNCTIONS FOR JUDGMENTS OF
 CONSISTENT SUBJECTS IN 4 ATTRIBUTE GROUP

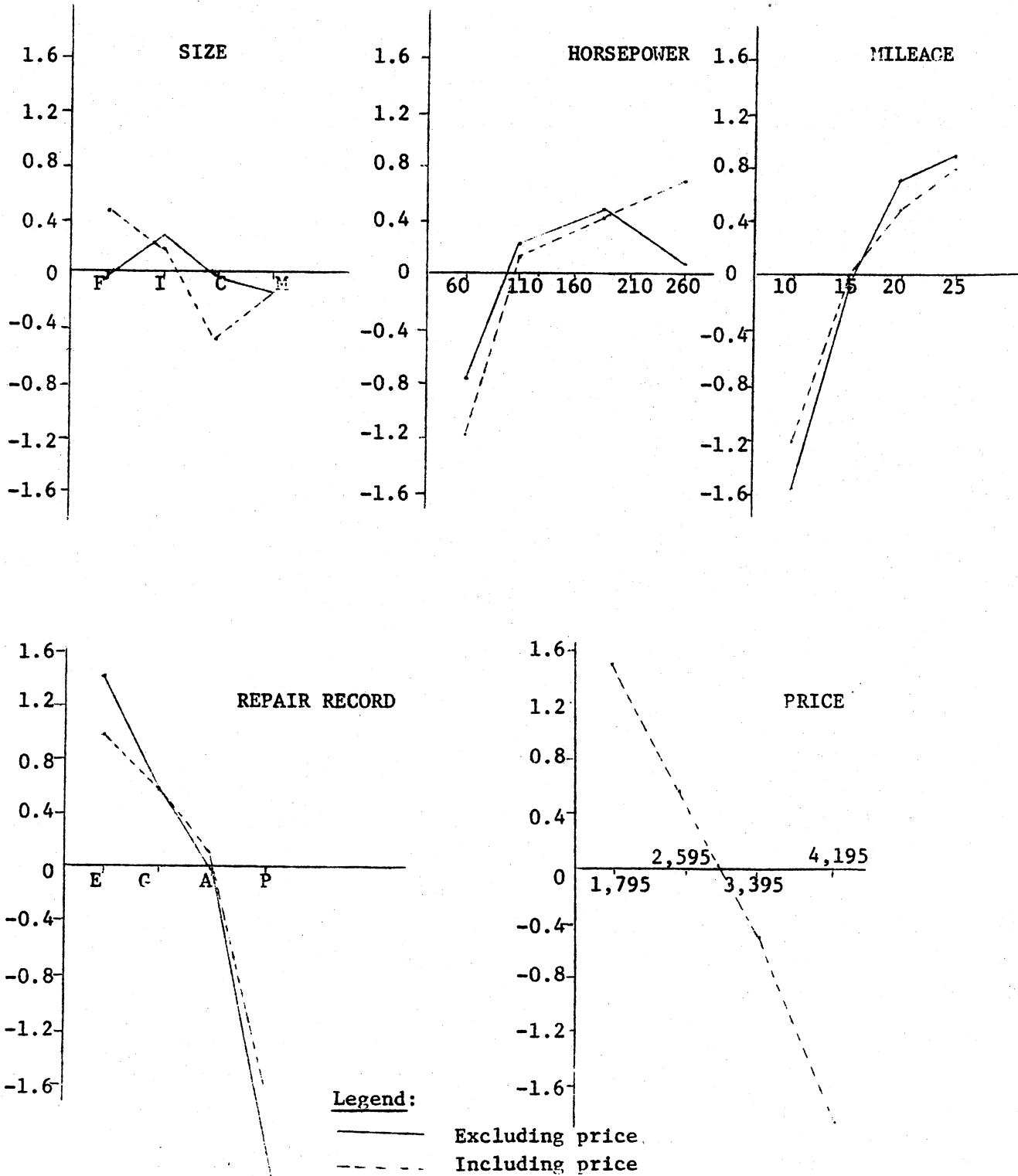


FIGURE 5

PARTWORTH FUNCTIONS FOR JUDGMENTS (EXCLUDING AND INCLUDING PRICE) OF CONSISTENT SUBJECTS IN 5 ATTRIBUTE GROUP

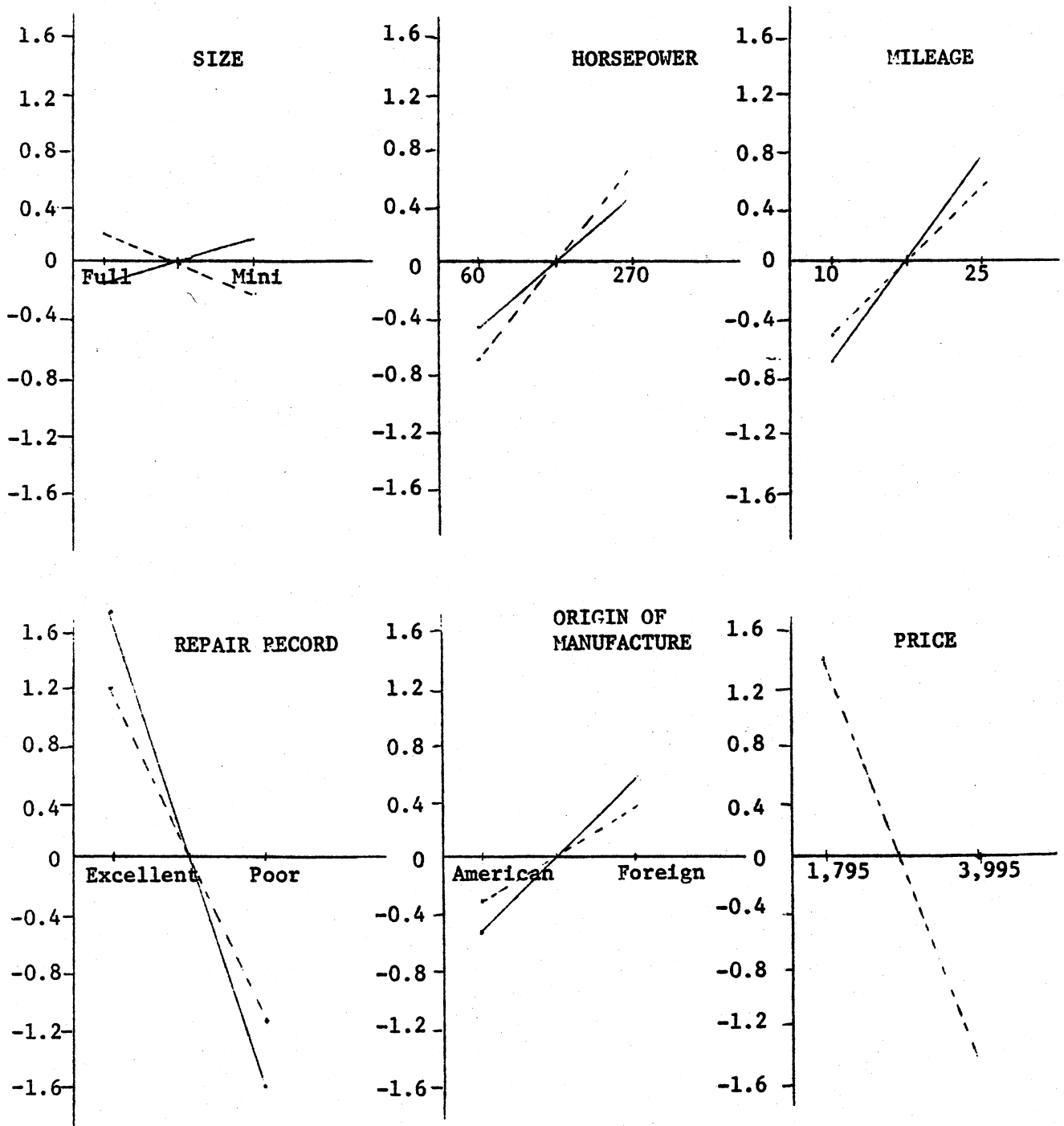
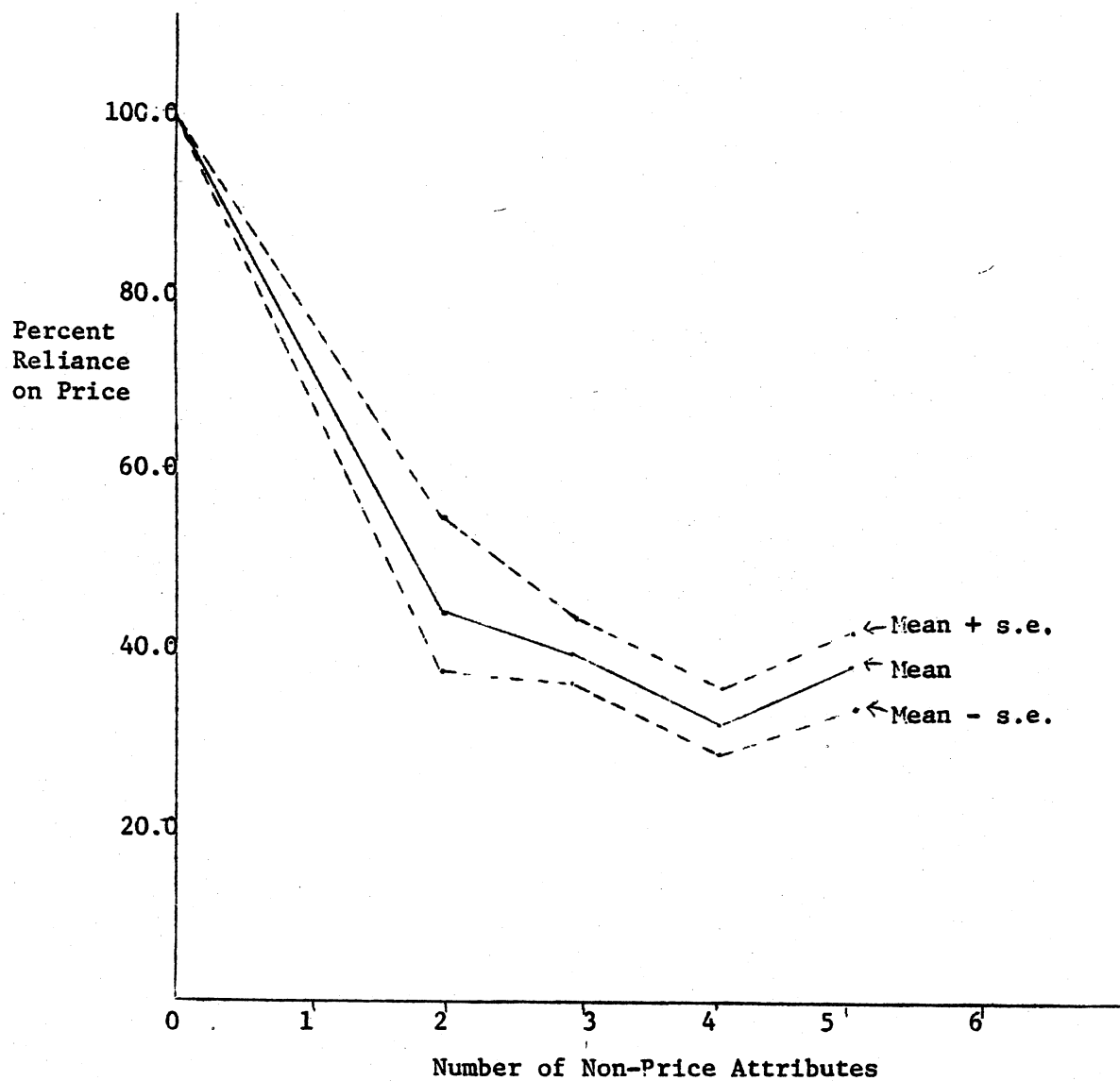


TABLE 7
 PERCENT CONTRIBUTIONS OF VARIOUS NON-PRICE AND PRICE
 ATTRIBUTES BY GROUP FOR CONSISTENT SUBJECTS

Number of non-price attributes	Sample size	Presence or absence of price information	Statistics ^a	Percent Contribution of					
				Price (P)	Size (A)	Horse-power (B)	Mile-age (C)	Repair record (D)	Origin of Manufacture (E)
2	14	Absent	Mean		31.2	68.8			
			s.e.		8.2	8.2			
		Present	Mean	45.8	11.5	42.7			
			s.e.	8.0	3.2	7.7			
3	22	Absent	Mean		13.0	38.3	48.7		
			s.e.		3.0	5.1	5.5		
		Present	Mean	40.8	9.6	33.0	16.6		
			s.e.	5.2	1.9	4.4	4.3		
4	21	Absent	Mean		5.9	15.7	28.2	50.2	
			s.e.		1.2	4.6	3.8	5.5	
		Present	Mean	32.0	8.1	16.7	15.9	27.3	
			s.e.	4.2	1.8	2.8	2.6	4.1	
5	20	Absent	Mean		5.2	8.3	15.1	62.8	8.6
			s.e.		1.6	2.5	4.0	6.7	3.1
		Present	Mean	38.3	5.5	15.0	7.1	30.9	3.2
			s.e.	5.1	1.7	5.4	2.1	6.1	0.9

^aThe entries are the mean percent of total contribution and its standard error.

FIGURE 6
PERCENT RELIANCE ON PRICE IN JUDGEMENTS OF
CONSISTENT SUBJECTS BY NUMBER OF NON-PRICE ATTRIBUTES



Conclusions

While the above analysis is indicative of ways in which individuals process attribute information on alternative stimuli, the results need to be replicated with more representative samples and with real brand names to be of use in marketing planning and policy formulation. However, two very tentative implications may be noted. First, from the public policy point of view, this study points out some directions as to the number of non-price attributes on which information should be disseminated to consumers enabling a better purchase decision. Furthermore, more attention and research seems necessary as to which non-price attributes should be emphasized. A similar implication follows for promotional planning by marketers of brands.

Footnotes

1. The author expresses his thanks to Messers. Michael Harper and Leonard Fertuck for their assistance in data collection and analysis and the Office of Sponsored Research of Cornell University for partial financial support.
2. MONANOVA program analyses data from a factorial experiment (fractional or otherwise) and finds the transformation which reduces the interaction as much as possible. It finds a monotone transformation of the data so as to achieve the highest possible percentage of variance accounted for by main effects. The program can handle any number of missing observations.
3. Some description of the sample characteristics may be in order here. Although not reported in the main paper, data were collected on two background characteristics, namely, annual income and years of driving experience. The means and standard deviations of these for the experimental groups are shown below:

<u>Group</u>	<u>Income (in \$000's)</u>		<u>Years of Driving</u>	
	<u>Mean</u>	<u>S.D.</u>	<u>Mean</u>	<u>S.D.</u>
2 Attributes	7.62	2.87	9.79	9.32
3 Attributes	7.89	4.75	9.64	7.24
4 Attributes	7.09	3.25	8.52	8.75
5 Attributes	6.42	1.62	7.26	3.15

In addition, data on a set of 12 attribute-interest-opinion questions tapping the individual's attitude toward cars, pollution and car maintenance were collected. These data (after transforming into factor scores) were employed to discriminate the four groups of subjects with the following results:

<u>Actual</u>	<u>Predicted by Discriminant Functions</u>			
	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>
2	8	3	8	5
3	16	6	5	9
4	8	5	4	10
5	8	4	3	16

The rate of correct classification was 34 percent.

References

- Coombs, Clyde H., Dawes, Robyn M. & Tversky, Amos. Mathematical Psychology: An Elementary Introduction. Englewood Cliffs, N. J.: Prentice-Hall, 1970.
- Coombs, C. H. A Theory of Data, New York: John Wiley and Sons, 1964.
- Gabor, A. & Granger, C. W. J. Prices as an Indicator of Quality: Report on an Inquiry. Economica, 1966, 33, 43-70.
- Gardener, David W. Is There a Generalized Price-Quality Relationship? Journal of Marketing Research, 1971, 8, 241-3.
- Green, Paul E., Carmone, Frank J., & Wind, Yoram. Consumer Evaluation of Discount Cards: A Conceptual Model and Experimental Test. Working Paper, University of Pennsylvania, February, 1971.
- Green, Paul E. & Rao, Vithala R. Conjoint Measurement for Quantifying Judgmental Data. Journal of Marketing Research, 1971, 8, 355-63.
- Howard John A. & Sheth, Jagdish N. The Theory of Buyer Behavior. New York: John Wiley and Sons, 1969.
- Jacoby, Jacob, Olson, Jerry C. & Haddock, Rafael A. Price, Brand Name, and Product Composition Characteristics as Determinants of Perceived Quality. Purdue Papers in Consumer Psychology, Paper No. 111, 1970.
- Kruskal, Joseph B. Analysis of Factorial Experiments by Estimating Monotone Transformations of the Data. Journal of the Royal Statistical Society, Series B, 1965, 27, 251-63.
- Leavitt, Harold J. A Note on Some Experimental Findings About The Meaning of Price. Journal of Business, 1954, 28, 205-10.
- Luce, R. Duncan. Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement. Journal of Mathematical Psychology, 1964, 1, 1-27.
- McConnell, J. Douglas. An Experimental Examination of the Price-Quality Relationship. Journal of Business, 1968a, 40, 439-44.
- McConnell, J. Douglas. The Price-Quality Relationship in an Experimental Setting. Journal of Marketing Research, 1968b, 5, 300-3.
- Peterson, Robert A. The Price-Perceived Quality Relationship: Experimental Evidence. Journal of Marketing Research, 1970, 7, 528-8.
- Rao, V. R. Saliency of Price in the Perception of Product Quality: A Multidimensional Measurement Approach. In Niel H. Bordon, Jr. (ed.), Proceedings of the 1971 Fall Conference of American Marketing Association, Chicago, Illinois (1972).
- Roskam, E. E. Metric Analysis of Ordinal Data in Psychology. Van Voorschoten, 1968.
- Scitovsky, Tibor. Some Consequences of the Habit of Judging Quality by Price. Review of Economic Studies, 1944-45, 12, 100-05.
- Stafford, James E. & Enis, Ben M. The Price-Quality Relationship: An Extension. Journal of Marketing Research, 1969, 6, 456-8.
- Tull, S. S., Boring R. A. & Gonsoir M. H. A Note on the Relationship of Price and Imputed Quality. Journal of Business, 1964, 38, 186-91.
- Tversky, Amos. A General Theory of Polynomial Conjoint Measurement. Journal of Mathematical Psychology, 1967a, 4, 1-20.
- Tversky, Amos. Additivity, Utility and Subjective Probability. Journal of Mathematical Psychology, 1967b, 4, 175-201.

and economic significance and quality differences for the two products under study (slacks and gasoline) as well as for 38 "filler" products; (2) highest and lowest prices considered "reasonable" for the 40 products; and (3) willingness to buy one of two alternatives, at specified prices, of slacks and gasoline.

Price limits were obtained with the question: "An average price is given for each product or service listed below. The average price was based on a sampling of stores and relevant businesses in this area. For each product, indicate the highest price that it would seem reasonable for a merchant to charge for that product. Similarly, also write the lowest price you think is reasonable. In other words, how high could the price be before you would think the product was overpriced? How low could the price be before you would find the low cost unbelievable?" Subjects wrote their responses in blanks provided. The average mean price for gasoline was \$.38 a gallon, for a pair of slacks \$9.40.

Price differential and base were manipulated for slacks and gasoline. The basic question for slacks was: "Suppose that you have to choose a new pair of slacks on the basis of the information given in the questions below. This is all the information you have. Which pair would you select?" Two price alternatives were presented: "Pair A sells for \$___; Pair B sells for \$___." Subjects responded on a nine-point scale containing nine equally spaced gradations labeled according to willingness to buy each of the alternatives. Endpoints were "definitely buy A" and "definitely buy B."

Nine pairs of prices, representing three differentials at each of three bases, were presented for each product. According to pretests, these price differentials assured prices beyond as well as within both upper and lower price limits. They also provided equal price intervals as manipulations. Stimulus pairs appear in Table 1.

Table 1
Price Manipulations

GASOLINE			
Differential	1	2	3
	\$.02	\$.06	\$.10
Base 1 \$.18	.19/.17	.21/.15	.23/.13
2 \$.28	.29/.27	.31/.25	.33/.23
3 \$.38	.39/.37	.41/.35	.43/.33
SLACKS			
Differential	1	2	3
	\$1.00	\$3.00	\$5.00
Base 1 \$ 4.00	4.50/3.50	5.50/2.50	6.50/1.50
2 \$ 9.00	9.50/8.50	10.50/7.50	11.50/6.50
3 \$14.00	14.50/13.50	15.50/12.50	16.50/11.50

preference for the lower-priced of two alternatives increased with price base, but the rate of increase and differential were positively related. Nystrom (1970) reported a tendency for consumers to generalize between the price of any particular item and the generic price associated with a product assortment. Despite the apparent effectiveness of subtle price variations, researchers frequently select prices to represent approximately equal divisions of the range of market prices (Lambert, 1970; McConnell, 1968; Valenzi & Andrews, 1971).

Responsiveness to subtle price changes is basic to latitude of acceptance and rejection for price. Latitude of acceptance refers to the range of stimulus values judged acceptable by an individual, while latitude of rejection encompasses the range of objectionable stimuli (Sherif & Sherif, 1969). However, despite the fact that techniques for reliably measuring these latitudes have been described by Gabor and Granger (1966) and Monroe and Venkatesan (1969, 1970), they have not been used to clarify the function of price in consumer decisions.

A consumer should be more receptive to a price which falls within his latitude of acceptance. Because individuals differ in the particular prices that represent upper and lower limits for acceptability, price will interact with limits. An individual with a relatively low lower limit is less likely to find a price overly low than the individual with a relatively high lower limit. Similarly, a relatively low upper limit increases the likelihood that a price will be perceived as unacceptably high. Because some prices are unacceptable to almost all consumers for certain products, interaction between price and individual limits is most likely at the extremes of the product price ranges. Sherif (1963) presented prices that exceeded individuals' upper price limits. Although all subjects tended to shift their latitudes of acceptance upward, the degree of shift depended upon initial upper limits.

Because both base and differential contribute to the total price, both price components should be affected by price limits. At a price base near either limit, large rather than small differentials are more likely to produce a price which falls beyond the limit. However, the extent of differential price change is unidimensionally restricted. Theoretically, prices can increase but not decrease indefinitely. The range of price differentials which produce plausible prices appears to be more restricted near lower than near upper price limits. In this case, small differentials are more important at the lower than upper limits. Small differentials contribute more toward the possibility of an unacceptable price at low price bases. Thus, small and large differentials should evoke greater behavior change at low and high price bases, respectively. According to Kamen and Toman (1970), large differentials evoked greater change in inclination to buy regardless of base. If the price stimuli tended to fall near and beyond the upper limit, this interaction would be expected.

Given a small differential, low price base should affect behavior more than a high price base. At a high differential the reverse would be true. Kamen and Toman (1970) found that price base was more important in determining preference if the differential was large rather than small. Failure to include prices perceived as unacceptably low would also yield these results.

Without attention to price limits, interpreting the nature of relationships between price and evaluation of products is difficult. As noted, price manipulations often represent the market range. Inclusion of unacceptably low prices is unlikely. Given market prices, the lower price limit tends to approximate the lowest price stimulus presented (Gabor & Granger, 1966). Over-representation of unacceptable high prices would elicit a negative relationship between price and perceived quality or preference without clearly detecting the relationship for acceptable prices alone. Gardner (1971) restricted prices to an acceptable range and found a positive relationship between price and likelihood of purchase. Lambert (1970) did not restrict prices and detected a negative relationship for the same product. Although suggestive, this difference was obtained without controlling for discrepancies between experiments other than the price range used.

Two experiments were conducted to explore the simple and interactive effects of price base, differential, and latitudes of acceptance on product evaluation and selection. The first used a single measure of price limits for two products. In the second experiment, several measures of price limits were examined. The possible influence of demand characteristics on limit measures was also evaluated. Both experiments manipulated price base and differential.

Given a situation in which an individual chooses between two differently priced products which are otherwise identical, the following relationships are hypothesized. (1) Perceived quality and willingness to buy the less costly alternative are positively related to both price level and price differential. Perceived quality is frequently inferred from willingness to buy (Leavitt, 1954; Tull et al., 1964), although the two responses are clearly separable. This hypothesis assumes that price negatively affects selection of a relatively costly product, due to desire for economy, but carries connotations of quality nonetheless. (2) Price base and individual price limits interact. Perceived quality is relatively low when price falls below the lower price limit, or exceeds the upper limit. Willingness to buy is relatively low when price is unacceptably low or high and is higher for acceptable prices. (3) Price base and differential interact. Specifically, the positive relationship between perceived quality or preference for the less costly alternative and price differential is attenuated at a low price base. Small differentials tend to be more effective at a low than at a high price base, while large differentials tend toward less effectiveness at a low base. This hypothesis combines the effect of price limits for all individuals with the restriction on effective price changes at low price levels.

Experiment 1

Method

The subjects were 44 college students who volunteered to participate in a psychological experiment. All had purchased the products involved in the experiment at least 15 times. Twenty females and 24 males participated.

Written questionnaires were administered in a single sitting of approximately one hour. The experimenter introduced the questionnaire as a survey of consumer attitudes toward various aspects of purchase decisions. Subjects provided three basic types of information: (1) ratings of social

Subjects also evaluated their chosen alternative on four 9-point semantic differential scales with endpoints of high quality-low quality, economical-uneconomical, fashionable-unfashionable, and good workmanship-poor workmanship. Perceived product quality is frequently posited as a mediator of product preference. Product attributes with financial, social, and functional implications may affect both perceived quality and preference.

Filler items were used to divide responses to price manipulations into several portions. Price manipulations were randomly selected for each portion. Order of portions was randomized across subjects.

Analysis

Measures for the five dependent variables were used in a combination of univariate and multivariate analyses of variance (Hummel & Sligo, 1971). A multivariate analysis of variance for each product examined the dependent measures as a function of price base and differential in a 3 x 3 design.

Means for highest and lowest reasonable prices were calculated for each product to represent the mean upper and lower limits, respectively. Each individual was classified according to whether each of his limits was above or below the group mean.

Following the multivariate analysis, univariate analyses of variance for each product incorporated four factors in a 3 x 3 x 2 x 2 design. Three levels of price differential and price base were within-subject factors. Classification above or below the upper and lower limits provided two levels for two between-subject factors.

Results

Means, standard deviations, and intercorrelations of the five dependent measures appear in Table 2. Perceived workmanship, fashionableness, and quality intercorrelated highly. These three variables tended to correlate negatively with perceived economy. Willingness to buy was not significantly correlated with any other dependent measure for either product.

Given the low correlations among several of the dependent variables, a 3 (differential) x 3 (base) multivariate analysis of variance (Winer, 1971, pp. 232-240) was carried out for each product to test the hypothesis of equal population mean vectors. The hypothesis was rejected for base ($F = 11.5033$, $df = 5,39$, $p < .001$) and differential-base interaction ($F = 3.9128$, $df = 20,558.14$, $p < .001$) for slacks. Similarly, the effects of base ($F = 38.0891$, $df = 5,39$, $p < .001$) and interaction between differential and base ($F = 5.410$, $df = 20,558.14$, $p < .001$) were significant for gasoline. Base and interactive effects produced unequal treatment means for at least one dependent variable. Price differential effect was not significant for either product.

Univariate analyses of variance were conducted for each dependent variable. This procedure is not overly conservative and assures a consistent experiment-wise error rate (Hummel & Sligo, 1971). Analysis of variance for repeated measures is appropriate if two assumptions are met: homogeneity of the variance-covariance matrices and symmetry of the pooled variance-covariance matrix. In each analysis the first assumption was tested with Hartley's F_{max} value. In no case was the hypothesis of homogeneity of

Table 2

Means and Intercorrelations^a for Five Dependent Measures (n = 44)

Measures	Slacks		Gasoline		Intercorrelations ^b				
	\bar{X}	S.D.	\bar{X}	S.D.	1	2	3	4	5
1 Willingness to buy costlier item (9 = high)	2.876	1.82	3.679	1.72	-----	.106	.024	.034	.099
2 Perceived quality (9 = low)	4.235	1.85	3.836	1.81	-.221	-----	.397**	.747**	.804**
3 Perceived economy (9 = high)	3.636	2.39	4.167	2.34	.258	-.333*	-----	-.380**	-.387**
4 Perceived fashionableness (9 = low)	4.715	1.42	3.730	1.62	-.178	.665**	-.376*	-----	.814**
5 Perceived workmanship (9 = low)	4.242	1.68	3.707	1.71	-.208	.852**	-.348*	.714**	-----

^aCorrelations are Pearson, parametric.^bCorrelations below diagonal are for slacks, above diagonal for gasoline.* $P < .05$ ** $P < .01$

variance rejected. The second assumption was not tested. Consequently, the degrees of freedom used to evaluate F values are those for the approximation procedure of Greenhouse and Geisser (1959).

Products were included as one factor in $3 \times 3 \times 2$ analyses of variance which contained base, differential, and product effects as within-subject factors. There were no main nor interactive effects involving product differences for any dependent variable. Base, differential, and differential-base interaction produced very similar effects for four dependent measures: willingness to buy, perceived quality, fashionableness, and workmanship. Accordingly, results will be reported for only one of these variables, willingness to buy, and the fifth variable, perceived economy. Results involving perceived economy did not approximate those of any other dependent variable.

Table 3 presents the results of the $3 \times 3 \times 2$ analysis for willingness to buy. Price base ($F = 54.9057$, $df = 1,43$, $p < .001$) and differential-base interaction ($F = 10.0368$, $df = 1,43$, $p < .005$) were significant. Inter-mean differences for price base ($\bar{X}_1 = 4.3826$, $\bar{X}_2 = 3.6742$, $\bar{X}_3 = 2.6970$) were significant beyond the .05 level in Newman-Keuls analysis. Willingness to buy the less costly alternative increased with price base.

Table 3
Analysis of Variance of Willingness to Buy
for Product and Price Differences

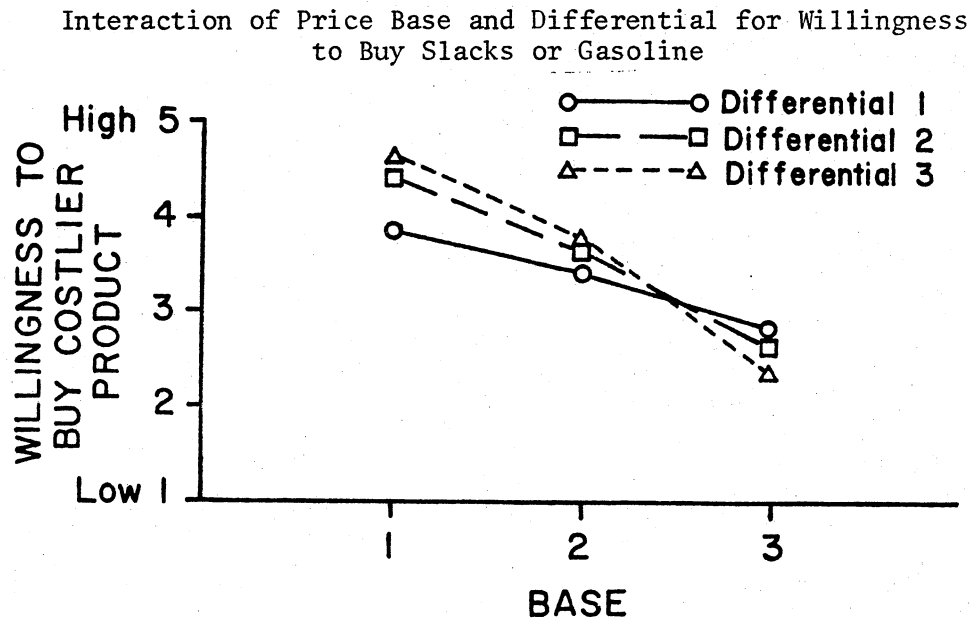
Source			
A (Differential)	2	2.660	1.4654
Error	86	1.815	
B (Base)	2	189.115	54.9057**
Error	86	3.444	
C (Product)	1	.102	.0086
Error	43	11.885	
A X B	4	8.369	10.0368*
Error	172	.834	
A X C	2	.216	.1128
Error	86	1.914	
B X C	2	.527	.2062
Error	86	2.554	
A X B X C	4	1.924	2.1307
Error	172	.668	

** $p < .001$, $df = 1,43$

* $p < .005$, $df = 1,43$

Differential-base interaction appears in Figure 1. At the low price base, willingness to buy the less expensive alternative decreased with differential ($\bar{X}_1 = 4.6818$, $\bar{X}_2 = 4.4773$, $\bar{X}_3 = 3.9886$). All inter-mean comparisons for differential at this base were significant ($p < .01$) in Newman-Keuls tests. At the medium base, means did not differ with differential. At the high price base, means differed insignificantly for small and middle differentials. At the large differential, willingness to buy the less expensive alternative was significantly less ($p < .05$) than for the smaller differentials.

Figure 1



Results of the 3 (differential) x 3 (base) x 2 (product) analysis of variance for perceived economy appear in Table 4. Price base and differential main effects were significant (respectively, $F = 12.7936$, $df = 1,43$, $p < .001$; $F = 20.1297$, $df = 1,43$, $p < .001$). Newman-Keuls analysis indicated significantly ($p < .05$) greater economy at the small differential relative to either larger differential. All inter-mean comparisons for base main effects were significant ($p < .01$) in Newman-Keuls tests, with greatest perceived economy at the medium base ($\bar{X}_1 = 4.1212$, $\bar{X}_2 = 5.3447$, $\bar{X}_3 = 4.5038$).

Interaction between base and differential was significant ($F = 27.7510$, $df = 1,43$, $p < .001$). Perceived economy decreased with differential at the low base (see Figure 2) and increased with differential at medium and high bases. Mean economy did not differ in Newman-Keuls tests for middle and large differentials at the medium base. All other inter-mean comparisons were significant ($p < .01$).

Evaluation of price limit effects involved separate analysis of variance for each product. Each subject was classified according to whether his highest and lowest reasonable prices fell above or below the group mean for that product. For gasoline the means for high and low reasonable prices \$.418 (s.d. = .0336) and \$.284 (s.d. = .0519), respectively. High and low group means for slacks were \$16.950 (s.d. = 4.918) and \$6.617 (s.d. = 1.990). Upper and lower limits were unrelated for slacks ($r = .132$, n.s.) and negatively related for gasoline ($r = .403$, $p < .01$).

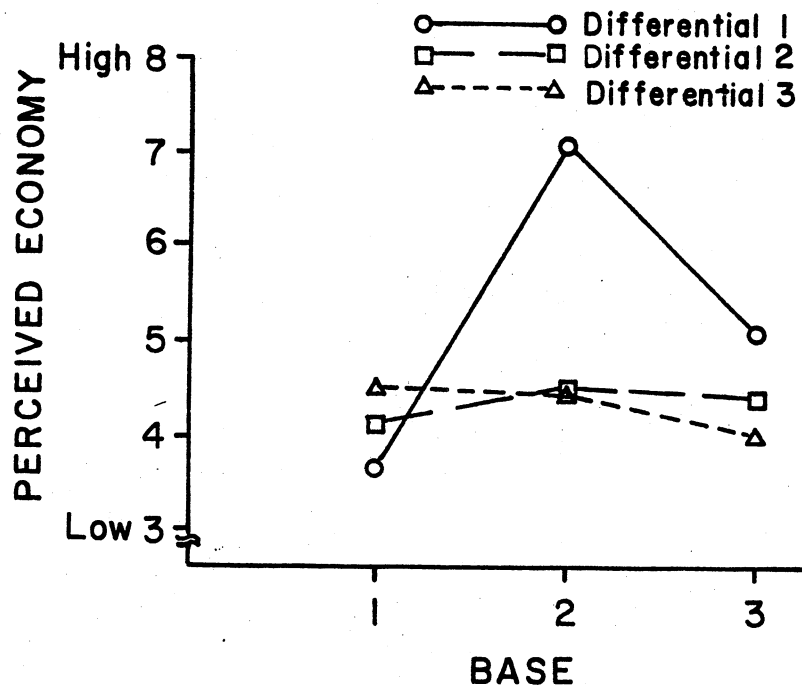
Table 4
 Analysis of Variance of Perceived Economy
 for Product and Price Differences

Source	df	MS	F
A (Differential)	2	73.338	12.7936**
Error	86	5.732	
B (Base)	2	10.418	20.1297**
Error	86	5.138	
C (Product)	1	1.293	.0687
Error	43	18.823	
A X B	4	79.166	27.7510**
Error	172	2.853	
A X C	2	.384	.0959
Error	86	4.003	
B X C	2	.783	.1150
Error	86	6.796	
A X B X C	4	9.151	1.2045
Error	172	3.946	

** $p < .001$, $df = 1,43$

Figure 2

Interaction of Price Base and Differential for
 Perceived Economy of Slacks or Gasoline



Analysis of variance of willingness to buy for the 2 (upper limit) x 2 (lower limit) x 3 (differential) x 3 (base) design yielded results very similar to analyses which excluded price limits. Base main effects and base-differential interaction were significant ($p < .05$) for both products. The pattern of interactive effects approximated that in the previous analyses. Interactions between lower limit and base for gasoline, and between upper limit and base for slacks, were marginally ($p < .10$) significant. For gasoline, the lower limit main effect was significant ($F = 6.1226$, $df = 1,38$, $p < .05$). Willingness to buy the less expensive alternative was greater for individuals with lower limits falling below the group mean.

With perceived economy of slacks alternatives as the dependent variable, analysis of variance of the 2 x 2 x 3 x 3 design yielded significant main effects for lower limit ($F = 11.8576$, $df = 1,40$, $p < .005$), price differential ($F = 6.5176$, $df = 1,40$, $p < .001$), and base-differential interaction ($F = 6.4354$, $df = 1,40$, $p < .05$). Perceived economy was less for individuals with a lower limit below the group mean ($\bar{X}_1 = 5.4773$, $\bar{X}_2 = 4.3605$). Newman-Keuls tests indicated significantly ($p < .05$) greater economy at the small differential than at middle or large differentials, which did not differ significantly. Base main effect and differential-base interaction corresponded closely to the cross-product results.

In Table 5 results of the 2 x 2 x 3 x 3 analysis of variance for perceived economy with gasoline as the product are presented. Significant base main effect ($F = 4.4516$, $df = 1,40$, $p < .05$) and differential-base interaction ($F = 21.5931$, $df = 1,40$, $p < .001$) were very similar to the cross-product results. The significant differential main effect ($F = 11.2736$, $df = 1,40$, $p < .05$) reflected decreases in perceived economy with increases in differential ($\bar{X}_1 = 4.9156$, $\bar{X}_2 = 4.1344$, $\bar{X}_3 = 3.7531$). All inter-mean comparisons for differential were significant ($p < .01$) in Newman-Keuls tests.

Three interactions among price factors and price limits were at least marginally significant: upper limit-differential ($F = 3.074$, $df = 1,40$, $p < .10$), both limits-differential ($F = 3.1553$, $df = 1,40$, $p < .10$), and lower limit-differential-base ($F = 5.2694$, $df = 1,40$, $p < .05$). The last three-factor interaction appears in Figure 3. In the low-lower limit condition, Newman-Keuls test indicated that perceived economy was significantly ($p < .01$) greater for the small differential than either of the larger differentials, which differ insignificantly. At the medium price base, economy was highest for the small differential, decreased at the large differential, and was lowest at the middle differential. All inter-mean comparisons at the medium price base were significant ($p < .01$). At the low base, mean economy was significantly ($p < .05$) greater at the middle differential than at large or small differentials.

At the low price base, within the high lower-limit condition, perceived economy was lowest for the large differential and highest for the middle differential. At the medium price base, perceived economy decreased as differential increased. Means for differentials at each base differed beyond the .01 level in Newman-Keuls analysis. At the lowest price base, mean economy differed insignificantly between middle and large differentials and decreased significantly ($p < .01$) at small differential.

Table 5

Analysis of Variance of Perceived Economy For Base
and Price Factors Within One Product: Gasoline

Source	df	MS	F
Between subjects:			
A (Upper limit)	1	.252	.0116
B (Lower limit)	1	47.799	2.2096
A X B	1	22.016	1.0136
Error	40	21.721	
Within subjects:			
C (Differential)	2	43.502	11.2736**
A X C	2	11.862	3.0740*
B X C	2	5.668	1.4690
A X B X C	2	12.175	3.1553*
Error	80	3.859	
D (Base)	2	30.111	4.4516**
A X D	2	1.317	.1947
B X D	2	1.123	.1660
A X B X D	2	3.108	.4594
Error	80	6.764	
C X D	4	57.240	21.5931***
A X C X D	4	2.564	.9678
B X C X D	4	13.969	5.2695**
A X B X C X D	4	.414	.1562
Error	160	2.651	

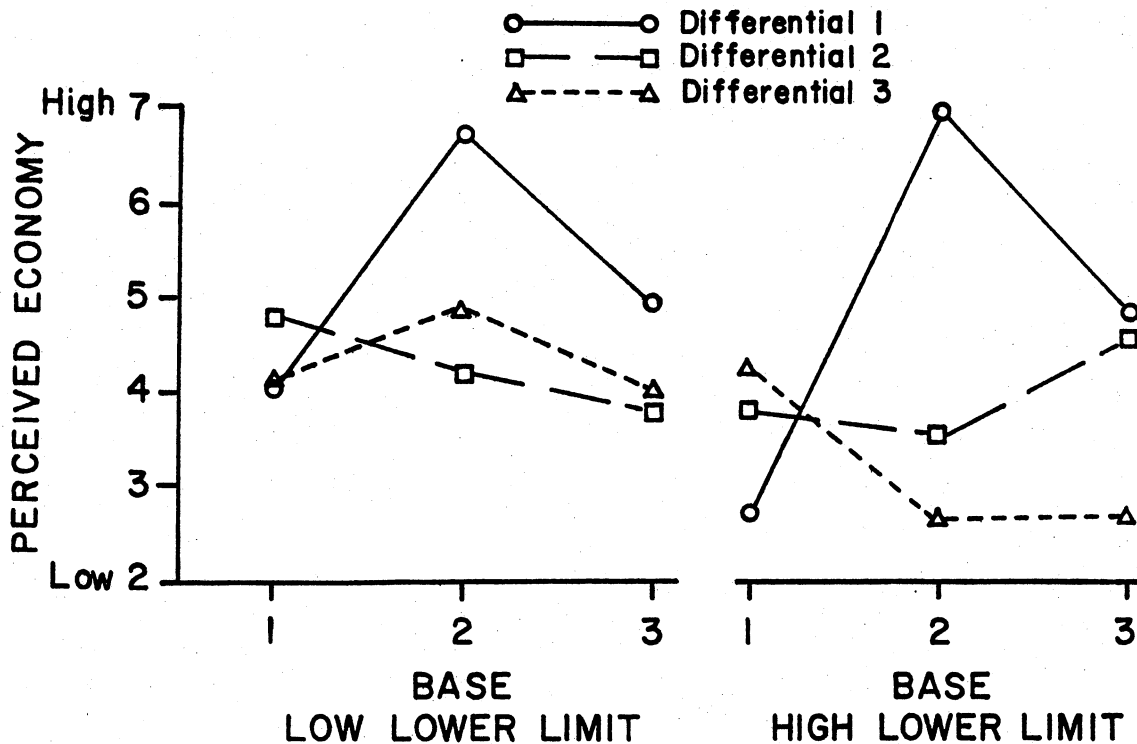
*** $p < .001$, $df = 1,40$

** $p < .05$, $df = 1,40$

* $p < .10$, $df = 1,40$

Figure 3

Interaction of Lower Limit, Price Base, and Price Differential for Perceived Economy of Gasoline



Experiment 2

Both price limits in Experiment 1 did not consistently interact with price manipulations as expected. It appears possible that the particular measure of price limits used may not have represented prices which were acceptable and unacceptable. Requesting prices that are "reasonable" may not elicit prices which the consumer considers acceptable for his own purchases. Consequently, a second experiment was conducted to examine several measures of price limits, including the own categories technique (Monroe & Venkatesan, 1970; Sherif & Hovland, 1961). Price base and differential were manipulated to evaluate the consistency of price effects for a third product and to examine the relationship of both price base and differential to price limits.

Method

Written questionnaires were completed by 44 of 50 female homemakers contacted by mail and living in Indianapolis, Indiana. Potential subjects were randomly selected from a listing of residential units in the city.

Each questionnaire was prefaced with a cover letter introducing the questions about "how you make buying decisions" and emphasizing the importance of each person's ideas. Comments about the questionnaire were solicited, and participants were offered \$1.00 upon receipt of their completed questionnaire.

Price base and differential were manipulated within the following basic question: "Imagine that you are in a store to buy a pair of shoes for everyday wear. You are considering two pairs that you like equally well in terms of size, style, and color. If the price of pair Y is \$___ and the price of pair Z is \$___, how likely is it that you would buy pair Y?" Four differentials (\$1, \$4, \$7, \$10) were used at each of three levels, \$7.50, \$19.50, and \$31.50. The seven-point response scale measured willingness to buy each alternative, with endpoints of "definitely buy pair Y" and "definitely buy pair Z." Letter designations were randomized across higher and lower-priced alternatives.

Price limits were measured in four ways. One measure used an adaptation of the own categories technique (Sherif, 1963). Subjects were asked to suppose that they were in a store to purchase shoes for everyday wear. Suitable color, style, and size were assumed. A list of prices was provided to help in selecting a pair of shoes, with these instructions: "We would like for you to first divide the price list into groups of prices. To help you start, we are providing you with two category labels for these price groups. They are: TOO EXPENSIVE TO BUY (TE) and TOO CHEAP TO BUY (TC). If you find any prices that you think are too expensive to buy, write TE to the right of each price. If you find any prices you think are too cheap to buy, write TC to the right of each price. Remember, these two labels are only provided as a starting point. You do not have to use them if you find no prices that belong in these two categories."

Half of the subjects were provided with optional endpoint labels of "very low in cost" and "very high in cost." Grouping prices within brackets was demonstrated with an example. Subjects were cautioned that the number of groups or prices in a group were unimportant, as long as prices were grouped "on the basis of which prices seem to belong together." Prices ranged from \$1.99 to \$35.99 in \$1 intervals.

Following the grouping task, subjects were asked to label one group "most acceptable" (MA), one group "most unacceptable" (MUA), and for any unlabeled groups considered acceptable (A) or unacceptable (UA), those respective labels.

Additional questions required the highest and lowest prices each subject remembered ever having paid for shoes and the highest and lowest reasonable prices. Subjects were also asked about the perceived purpose of the questionnaire, their ease in understanding the various tasks, and several demographic characteristics.

Results

Demand Characteristics

Four indices of price limits were compared for the "too cheap/too expensive" label condition and the "too low/high in cost" label condition. The mean price, range of prices, mean lower price and mean upper price for the range of prices constituting the lower (upper) limit was computed for each set of labels. Label differences did not significantly alter any of the measures. Labeling conditions also did not affect reported price paid previously, reasonable prices, or the prices labeled acceptable and unacceptable.

Comparison of Price Limit Measures

Means and intercorrelations in Table 6 represent four measures of price limits: prices paid previously, prices considered reasonable, prices labeled very low cost (cheap) or high cost (expensive), and prices labeled acceptable or unacceptable. Previously paid and reasonable prices corresponded closely at each limit, correlating positively at both lower ($r = .491$, $p < .001$) and upper ($r = .656$, $p < .001$) limits. MUA price correlated negatively with low price paid previously ($r = -.313$, $p < .05$) or considered reasonable ($r = -.470$, $p < .01$).

The negative relationship between MUA and reasonable or previously paid prices reflects the fact that 39 of the 44 respondents classified either high or low prices, but not both, as unacceptable. The mean MUA price was \$18.26, indicating that high prices tended to be unacceptable for most subjects. However, 14 subjects perceived a range of low prices as MUA. The mean MUA price for the group was \$3.57. That few individuals perceived both low and high prices as unacceptable contrasts with several studies (Gabor & Granger, 1964, 1966; Sherif, 1963).

Price Limits and Price Changes

The relationship of price limits to price base and differential was examined for two indices of limits, the own categories labeling and prices considered reasonable.

Splitting the sample at mean MUA price and mean MA price provided cell entry data for a $2 \times 2 \times 4 \times 3$ analysis of variance, with two levels of each limit, four price differentials, and three price bases (Table 7). Base and differential main effects and differential-base interaction were each significant ($df = 1,37$, $p < .05$). Also significant were the MA and MUA main effects (respectively, $F = 9.7611$, $df = 1,37$, $p < .005$; $F = 6.9725$, $df = 1,37$, $p < .05$). Interaction between MA and base, and interaction among MA, differential, and base were significant beyond the .05 ($df = 1,37$) level. MA-differential interaction was marginally significant ($p < .10$, $df = 1,37$).

Newman-Keuls analysis indicated significant ($p < .01$) decrease in willingness to buy the costlier alternative between the low and medium bases ($\bar{X}_1 = 3.987$, $\bar{X}_2 = 2.959$). Mean willingness at medium and high bases differed insignificantly for differentials one and two, the lower differentials, and for the higher differentials three and four, but decreased significantly ($p < .05$) between the smaller and larger differentials.

Interaction between differentials and bases appears in Figure 4. At the low base, means for willingness to buy at differentials, one, two, and four differed insignificantly in Newman-Keuls tests but indicated significantly ($p < .01$) less tendency to buy the costlier alternative than at differential three. At the medium base, no significant differences in means occurred. Means between differentials one and two, and between three and four, differed insignificantly in Newman-Keuls analysis for differential effects at the high base. Willingness to buy the costlier alternative at the larger differentials is significantly less ($p < .01$) than at the smaller differentials.

Main limit effects represent significantly greater willingness to buy the costlier alternative when either the MA or MUA price is relatively high. Newman-Keuls analysis of means in the MA-differential interaction indicated

TABLE 6

Means and Intercorrelations of Four Measures of Price Limits

Measures	Intercorrelations							
	1	2	3	4	5	6	7	8
1 Lowest price paid previously	----							
2 Lowest price considered reasonable	.491***	----						
3 Very low in cost/too cheap price	.078	.499***	----					
4 Mean, most acceptable prices	.371*	.259	.094	----				
5 Mean, most unacceptable prices	-.313*	-.446**	-.149	-.092	----			
6 Highest price paid previously	.263	.268	.026	.470**	-.336*	----		
7 Highest price considered reasonable	.169	.279	-.125	.388*	-.275	.656***	----	
8 Very high in cost/too expensive price	.159	-.047	.345*	.169	.334*	-.127	-.428**	----
\bar{X}	5.555	5.422	2.965	10.116	18.256	24.026	24.162	25.826
S.D.	5.448	3.921	1.688	5.606	13.104	7.519	13.982	9.036

*** $p < .001$ ** $p < .01$ * $p < .05$

TABLE 7

Analysis of Variance of Willingness to Buy for
Limit and Price Factors Within One Product: Shoes

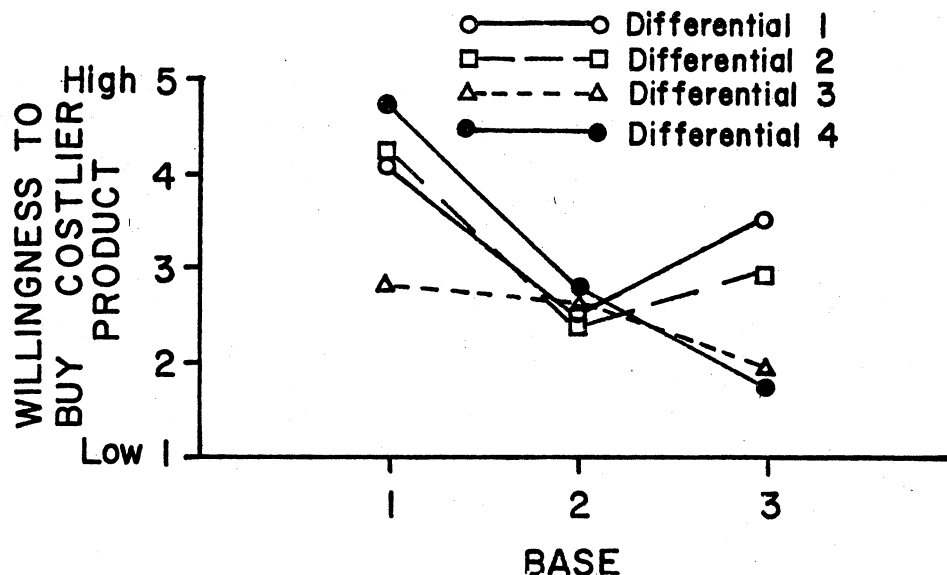
Source	df	MS	F
Between subjects:			
A (Most unacceptable prices)	1	21.854	6.9725*
B (Most acceptable prices)	1	30.595	9.7611**
A X B	1	.331	.1056
Error	37	3.134	
Within subjects:			
C (Differential)	3	19.700	7.2151*
A X C	3	4.202	1.5389
B X C	3	10.198	3.7348
A X B X C	3	1.348	.4936
Error	111	2.730	
D (Base)	2	99.994	25.4939**
A X D	2	2.914	.7430
B X D	2	16.866	4.3002
A X B X D	2	7.269	1.8534
Error	74	3.922	
C X D	6	18.808	6.9226*
A X C X D	6	2.631	.9883
B X C X D	6	13.314	4.9006*
A X B X C X D	6	6.445	2.3721
Error	222	2.717	

** $p < .005$, $df = 1,37$

* $p < .05$, $df = 1,37$

Figure 4

Interaction Between Price Base and Differential
For Willingness to Buy Shoes



significantly ($p < .01$) higher willingness to buy at differentials one and four for individuals with high MA prices. Mean willingness at differentials two and three did not differ according to MA limit. The significant MA-base interaction derives only from differences at the medium price base; mean willingness is greater for individuals with relatively high MA prices ($p < .05$ in Newman-Keuls analysis).

Although the four-factor interaction was marginally ($p < .10$) significant, simultaneous effects of both MA and MUA limits were of interest. Newman-Keuls tests clarified certain regularities in interactive effects. Variation in mean willingness to buy across differentials at the high price base was similar regardless of limit classification. Willingness to buy the costlier alternative tended to decrease with differential. Variation in differential affected willingness to buy at the medium base for a single group, those with relatively low MUA and MA prices. For this group, willingness to buy increased significantly from ($p < .05$) differentials one and two, which did not differ, to differential three, and again ($p < .05$) between differentials three and four. Limit effect was most notable at the low base. For low MUA price, willingness to buy the costlier alternative was greatest for differentials one and four, decreasing significantly ($p < .05$) at differentials two and three. For high MUA, the order of means according to decreasing willingness to buy varied with the MA limit. For high MA, the order was differential 1 = 2 = 4 > 3. For low MA, the order of decreasing willingness is differential 2 > 3 = 4 > 1.

The 2 X 2 X 4 X 3 analysis of variance with limits based on prices perceived as reasonable produced three significant effects: base ($F = 25.7938$, $df = 1,37$, $p < .001$); differential-base interaction ($F = 6.9262$, $df = 1,37$, $p < .05$); and the three-factor interaction of lower limit, differential, and base ($F = 5.4706$, $df = 1,37$, $p < .05$). Differential main effect, differential-lower limit interaction, and the four-factor

interaction were marginally ($p < .10$) significant. The base, differential, and differential-base effects paralleled those in the preceding analysis. Newman-Keuls analysis of means within the significant three-factor interaction revealed notable differences due to lower price limit only at the lower price base. At the medium price base, willingness to buy did not differ across differentials within each limit classification. At the high base, regardless of whether the individual had a relatively high or low lower limit, willingness to buy the costlier alternative decreased as differential increased. Decreases were significant ($p < .05$) between all differentials except three and four. For the low base, given individuals with relatively low lower limits, willingness to buy was greatest at differential two and significantly less ($p < .01$) at any of the other three differentials. The lowest as well as the two highest differentials thus increased willingness to buy the less costly alternative. If lower limits were relatively high, willingness to buy the less costly alternative increased significantly ($p < .05$) between differential one and two, and two and three, but was lowest at the largest differential.

Discussion

The positive relationship between price base and preference for less expensive alternatives predicted in Hypothesis 1 was found in both experiments, thus supporting earlier findings for low-priced, non-durable goods (Kamen & Toman, 1970; Lambert, 1970). Results of the current studies extend this relationship to two higher-priced, durable goods (slacks and shoes). Price base and perceived quality were also positively related, which confirms findings by Andrews and Valenzi (1971), Gardner (1971), McConnell (1968), and Valenzi and Andrews (1971).

Differential and willingness to buy were positively related as predicted in Hypothesis 1 for only one (shoes) of the three products studied. Because this product occurred alone in one experiment, either product-specific characteristics or experimental artifacts may be responsible for these differential effects. One explanation in terms of artifacts would involve relatively stronger differential manipulation in one experiment. Relative to mean price base, neither price differential intervals nor total range of prices within all intervals differed across products. Results may be due also to differences between samples in their attention to price. Female homemakers may be more sensitive to price differentials than students. Another possibility is increased sensitivity to differential for those who experienced four rather than three variations in differential.

Interaction between price base and differential occurred for all products, as posited in Hypothesis 3. However, the specific pattern of effects was only partially supportive. Interaction was most consistent at higher price bases. At the highest price base, preference for the less costly alternative was positively related to price differential. A higher absolute price elicited greater preference for the less costly product. At the low price base, differential usually altered dependent measures, but without any easily discernible pattern. Differential had little influence at the medium price base. Simple differential effects were consistently nonsignificant at this base. Price fluctuations about the medium base generally resulted in prices within individual price limits. Price effects appear to be relatively limited when both alternatives are clearly acceptable. Several measures of price sensitivity (Craig, Engel, & Talarzyk, 1971; Teach & Pessemier, 1966) assume this lower responsiveness for price changes within a certain range of the most acceptable price.

Price limits did not interact with price changes as predicted in Hypothesis 2. Marginally significant interactions occurred between price base and both lower limit (for gasoline) and upper limit (for slacks). The range of most acceptable prices interacted with both price level and differential for one product, shoes. At the low price base, price effect varied with price limits more than at the higher price bases. Particularly notable was the tendency of both extremely low and high differentials to produce greater preference for the more expensive alternative at the low base. This result occurred most clearly for individuals who found low rather than high prices most unacceptable, and for those who did not think extremely low prices were reasonable. In the experiments, a large differential at the low base was most likely to produce prices below the lower limit, leaving only the costlier item as acceptable. A small differential, like a larger one, conveys higher quality connotations for the more costly alternative (Experiment 1), but is more likely to function within the limits of acceptably low prices.

Despite similar effects for price base on both willingness to buy and perceived quality, the latter two variables were unrelated in this research. Quality is often assumed to be a mediator of product preference, either as one of the important criteria for selection or as a determinant of the salience of price (Leavitt, 1954; Tull et al., 1964). In contrast, the first experiment suggested that perceived quality and product preference constitute independent dimensions of response to price changes.

One of the more interesting findings about price base concerned another response dimension, perceived economy, which was unrelated to willingness to buy but covaried with perceived quality. Price base and perceived economy were not related in any simple negative manner. Perceived economy was greatest at the medium base, less at the high base, and least at the low base. Apparently, economy is not a direct reflection of financial expenditure. Evaluation of economy may encompass weighting price with associated costs over a longer period. A small present price may imply low immediate cost but higher long-term costs, perhaps through more frequent product replacement. In this case, the positive price-quality relationship may affect perceived economy.

Price differential altered perceived economy for the two products (gasoline, slacks) which were unaffected in terms of willingness to buy. Price differential and economy were negatively related. If perceived economy can be assumed to be more closely associated with price than willingness to buy, then the results suggest simply that the influence of a price change depends on how directly it is related to the dependent variable. The fluctuation in differential effect emphasizes the necessity for examining price relative to the evaluation of multiple product characteristics.

An unexpected result was the direct effect of price limits on responses. The most acceptable price was positively related to willingness to buy. Similarly, the highest reasonable price correlated positively with buying likelihood and economy, and negatively with quality. If price limits are surrogates for economic status, then the greater preference for costlier alternatives by those with greater resources is a relatively trivial finding. However, because directionality between limits and the dependent measures is not established in this study, higher acceptable prices may result from willingness or perceived economy. Numerous studies (e.g., Sherif, Sherif,

& Nebergall, 1965) have shown that involvement with stimuli alters the range of stimuli considered acceptable. Willingness to buy or perceived economy, as a reflection of product importance, may affect the absolute value as well as range of prices considered. Future research should recognize the possibility of mutual influence between price limits and evaluation or selection of products.

Interaction between price limits and changes in product price varied more with differences in most acceptable prices than most unacceptable prices. Measures of most unacceptable prices provide only a gross index of an individual's emphasis on low or high prices. Most unacceptable prices, by definition, include only one portion of the price continuum, usually the upper or lower extreme. Each extreme of the continuum is discriminated more finely by price limits based on prices paid previously, considered reasonable, or perceived as "too cheap" or "too expensive." Analyses using each of these four measures indicated that each was as useful as the distinction between most unacceptable and most acceptable prices in accounting for responses to price changes. If, as this study suggests, few persons spontaneously focus on both high and low prices as unacceptable, future researchers may find it advantageous to force distinctions at both extremes.

The basic findings of these studies can be briefly summarized as follows.

- (1) Preference for less costly product alternatives increased with price base. Perceived quality and price base were also positively related.
- (2) Price base and differential interacted to determine response to the three products. Given a relatively high price base, preference for the less expensive alternative, as well as perceived quality, increased with price differential.
- (3) Several logical definitions of price limits produced measures which did not intercorrelate highly. Demand characteristics in the experimental situation could not account for individual differences in price limits. Interaction between individual limits and price manipulations was weak and inconsistent.
- (4) Price differential and perceived economy were negatively related, while perceived economy followed an inverted-V pattern relative to price base. Interaction between differential and base involved high perceived economy for large differentials at a low price base, but low perceived economy for large differentials at a high price level.
- (5) Willingness to buy a costlier product alternative was unrelated to perceived quality, economy, fashionableness, and workmanship of the product.
- (6) In general, the studies reveal complex interactions between price changes that are small (differential) or large (base) relative to average product price. Both base and differential price changes alter perceived economy and quality, as well as preference, for product alternatives. Changes in price base appear to be more important, in that they show more consistent relationships across products than changes in price differential. Although perceived quality and willingness to buy vary monotonically with price base, perceived economy does not.
- (7) Individuals differ in their upper and lower price limits, as determined by a number of measures of limits. Limits do not apparently vary with certain demand characteristics of the experimental situation. Different measures of limits, however, are not highly related. The limits have some explanatory power, particularly when prices are relatively low. However, this efficacy for the limits is confounded with product differences and particular operational definitions for limits. Further research should concentrate on examining the possible generality of limit effects.

References

- Andrews, I. R., & Valenzi, E. R. Combining price, brand, and store cues to form an impression of product quality. Proceedings, 79th Annual American Psychological Association, 1971, 649-650.
- Craig, C. S., Engel, J. F., & Talarzyk, W. W. Consumer decision-making: On the importance of price. Proceedings, 2nd Annual Association for Consumer Research Convention, 1971, 243-255.
- Gabor, A., & Granger, C. W. J. Price sensitivity of the consumer. Journal of Advertising Research, 1964, 4, 40-44.
- Gabor, A., & Granger, C. W. J. Price as an indicator of quality: Report on an inquiry. Economica, 1966, 46, 43-70.
- Gardner, D. M. Is there a generalized price-quality relationship? Journal of Marketing Research, 1971, 8, 241-243.
- Greenhouse, S. W., & Geisser, S. On methods in the analysis of profile data. Psychometrika, 1959, 24, 95-112.
- Hummel, T. J., & Sligo, J. R. Empirical comparison of univariate and multivariate analysis of variance procedures. Psychological Bulletin, 1971, 76, 49-57.
- Jacoby, J., Olson, J., & Haddock, R. Price, brand name, and product composition characteristics as determinants of perceived quality. Journal of Applied Psychology, 1971, 55, 570-579.
- Kamen, J. M., & Toman, R. J. Psychophysics of prices. Journal of Marketing Research, 1970, 3, 27-35.
- Lambert, Z. V. Pricing and brand choice. AMA Fall Educators Conference, Boston, Massachusetts, 1970. (An abstract appears in David L. Sparks (Ed.), Broadening the concept of marketing. Chicago: American Marketing Association, 1970, p. 44.)
- Leavitt, H. J. A note on some experimental findings about the meaning of price. Journal of Business, 1954, 27, 205-210.
- McConnell, J. D. Effects of pricing on perception of product quality. Journal of Applied Psychology, 1968, 52, 331-334.
- Monroe, K. B., & Venkatesan, M. The concept of price limits and psychophysical measurement: A laboratory experiment. In P. R. McDonald (Ed.), Marketing involvement in society and the economy. Chicago: American Marketing Association, 1969, pp. 345-351.
- Monroe, K. B., & Venkatesan, M. The measurement of price thresholds: Psychophysics and latitudes of acceptance. AMA Fall Educators Conference, Boston, Massachusetts, 1970. (An abstract appears in David L. Sparks (Ed.), Broadening the concept of marketing. Chicago: American Marketing Association, 1970. p. 120.)
- Nystrom, H. Retail pricing. Stockholm: The Economic Research Institute at the Stockholm School of Economics, 1970.

- Orne, M. T. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. American Psychologist, 1962, 17, 776-783.
- Rao, V. Salience of price in the perception of product quality: A multi-dimensional measure approach. AMA Fall Educators Conference, Minneapolis, Minnesota, 1971.
- Sherif, C. W. Social categorization as a function of latitudes of acceptance and series range. Journal of Abnormal and Social Psychology, 1963, 67, 148-156.
- Sherif, C. W., Sherif, M., & Nebergall, R. E. Attitude and attitude change. Philadelphia: Saunders, 1965.
- Sherif, M., & Hovland, C. I. Social judgment: Assimilation and contrast effects in communication and attitude change. New Haven: Yale University Press, 1961.
- Sherif, M., & Sherif, C. W. Social psychology. New York: Harper and Row, 1969.
- Stafford, J. I., & Enis, B. M. The price-quality relationship: An extension. Journal of Marketing Research, 1969, 6, 456-458.
- Teach, P. D., & Pessemier, E. A. Pricing experiments, scaling consumer preferences, and predicting purchase behavior. In R. M. Hass (Ed.), Science, technology, and marketing. Chicago: American Marketing Association, 1966. pp. 541-560.
- Tull, D. S., Boring, R. A., & Gonsior, M. H. A note on the relationship of price and imputed quality. Journal of Business, 1964, 37, 186-191.
- Valenzi, E. R., & Andrews, I. R. Effect of price information on product quality ratings. Journal of Applied Psychology, 1971, 55, 87-91.
- Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1971.

CUE UTILIZATION IN THE QUALITY PERCEPTION PROCESS¹

Jerry C. Olson²
The Pennsylvania State University
and Jacob Jacoby³
Purdue University

From an information theoretic perspective, products and brands are conceived to consist of an array of cues (e.g., price, brand name, packaging, color, etc). Moreover, each cue provides a basis for developing various impressions of the product itself (cf. Cox, 1962; Jacoby, Olson, & Haddock, 1971). One impression of considerable importance in consumer behavior is the perceived quality of the product or brand. In addition to being of interest in its own right, quality perception seems to be strongly related to actual purchasing behavior, especially brand loyalty (cf. Jacoby, 1971).⁴

One cue to perceived product quality that has generated considerable recent experimental interest is price (Andrews & Valenzi, 1971; Enis & Stafford, 1969; Gardner, 1971; Jacoby, Olson, & Haddock, 1971; McConnell, 1968; Olson, 1972; Peterson, 1970; Rao, 1971; Szybillo & Jacoby, 1972; Valenzi & Andrews, 1971; and White, 1966). Other cues examined experimentally include brand image (Allison & Uhl, 1964; Jacoby, Olson, & Haddock, 1971), store image (Enis & Stafford, 1969; Szybillo & Jacoby, 1972), and actual composition differences (Jacoby, Olson, & Haddock, 1971; Szybillo & Jacoby, 1972; Valenzi & Andrews, 1971).

Examination of these studies leads to the conclusion they are basically exploratory and atheoretical. Although they are more realistic and sophisticated than the earlier studies of perceived product quality (cf. Laird, 1932; Leavitt, 1954; Tull, Boring, & Gonsior, 1964) in that they include more than a single quality cue, they nonetheless lack a theoretical framework and overall direction. It is the thesis of the present paper, to be developed later, that this lack of model and direction greatly hinders the study of cue utilization in the quality perception process.

One reason for the current state of affairs may be that all of the quality perception studies have utilized experimental designs which, although particularly useful for establishing cause and effect relationships, tend to drastically limit the number of cues that can be examined in combination. Witness the fact that the above-cited experiments have studied only one or two, occasionally three, and at the most four, cues simultaneously. Perhaps such limitations are understandable given the difficulty of interpreting higher order interactions when they occur.

Since only a few of the many possible cues can be incorporated into any one experiment, the critical question becomes which cues to examine. In answering, most researchers seem to have chosen quality cues for their investigations on the basis of intuition or specific interest in a particular cue (e.g., price). However, if one's purpose is to assess the relative effects of those cues which most strongly influence quality perceptions (hopefully with the intention of developing a comprehensive understanding of quality cue utilization), then it becomes imperative that the most salient cues first be identified so that they can be incorporated into one's experiment.

Accordingly, the present exploratory research attempted to answer the following question: Of those product attributes (or cues) that consumers take into consideration in deciding on a brand purchase, which cues are felt to have the greatest influence upon perceptions of brand quality? Since certain cues may be influential in some product categories but not in others, the present investigation sought to provide information concerning the strength of several potential quality cues across several product categories.

Method

Subjects

Sixty-nine housewives living in seven different socio-economic and geographic segments of Lafayette, Indiana, served as subjects.

Instrument

Data were collected using a self-administered questionnaire. Comprehensive lists of product-related attributes or potential quality cues were generated separately for five products selected to cover a wide range of prices and usage characteristics. In all, there were 15 product attributes for living room rug, 14 for hair dryer, 12 for ground coffee, 15 for shampoo, and 13 for aspirin tablets. Several of these attributes were product-specific (e.g., "type of fiber" for living room rugs); others were applicable to several products (e.g., "packaging" for aspirin, coffee, and shampoo); while others were common to all five products (e.g., "price" and "brand name"). Each product and its accompanying randomized list of product attributes were reproduced as a separate page in a six-page questionnaire. Excluding the first page of introductions, the order of the pages within the questionnaire was randomly determined for each subject.

Task

The housewives were asked to imagine that they were about to purchase each of the five products. For each product, the subject first checked those factors that she would take into consideration when making a decision about which brand of the product to purchase. Second, each subject rank ordered those factors she would consider in terms of their usefulness and importance as indicators of (or cues to) product quality.

Results

The data were analyzed separately for each of the five products. Two indices were computed for each product attribute: (a) the number of subjects who stated that they considered the attribute in making a purchase decision, and (b) the average ranking of each attribute (based on those subjects considering it) in terms of its perceived usefulness and importance as an indicator of or cue to product quality.

Parenthetically, these data are relevant to an issue regarding the number of important, desired product attributes used as evaluative criteria by consumers in choosing among alternative brands during the purchase decision process (cf. Engel, Kollat, & Blackwell, 1968, pp. 430-431). In the present research, the first cue index (i.e., frequency of consideration of a cue in a purchase decision) can be considered as a rough measure of "cue use as an evaluative criterion." The average number of product attributes these

consumers reported considering in each product category was: hair dryer = 5.97; living room rug = 7.17; ground coffee = 4.64; shampoo = 5.13; and aspirin tablets = 4.51. These data suggest that, for most consumers, several (rather than one or two) product attributes are considered during the evaluation process. Given this likelihood of multiple cue use, it becomes even more crucial to determine relative cue importance. It must be noted, however, that the data cited above are based upon self-reports rather than actual cue use and, therefore, may be inaccurate estimates of the "true" number of product attributes used as evaluative criteria.

Table 1 presents the four cues in each product category with the highest average ranks in terms of perceived importance or accuracy as indicators of product quality. One can easily detect a lack of commonality across product categories in terms of the specific attributes thought to be accurate indicators of product quality. This is especially evident for the most important cue in each product category. Accurate cues to quality seem to be highly product specific.

A brief digression is warranted at this point. Thus far, the results of this research provide yet another example of a common conclusion for consumer behavior research in general, and quality research in particular. That is, one's findings are likely to be specific to the type of product and/or consumer investigated. In such cases, generalizations beyond the product or consumers examined are of dubious validity. One strategy in such a situation is to develop concepts or constructs at a more abstract level than the observable data, which hopefully can explain confusing or conflicting empirical results and may enable one to generalize beyond the immediate situation.

After some thought regarding the lack of commonality in Table 1, such a potentially useful conceptual dimension was hypothesized. Note in Table 1 that the first or second most important cue to perceived quality generally was one which could be described as "intrinsic to the product"--i.e., a product attribute which cannot be changed or experimentally manipulated without also changing the physical characteristics of the product itself. In contrast, extrinsic cues are product-related attributes which are not a part of the physical product. For example, "nature of the fiber" (for living room rug), "taste" (for coffee), and "special ingredients" (for aspirin) are all integral, intrinsic components of the physical products. Conversely, product-related factors such as "guarantees" (for living room rug), "manufacturer's reputation" (for hair dryer), and "brand name" (for ground coffee) are extrinsic cues in that they are not a part of the physical product. In contrast to intrinsic cues, if extrinsic cues are experimentally manipulated the physical characteristics of the product need not necessarily change. The present study found that intrinsic, rather than extrinsic, cues were generally perceived to be the most accurate indicators of brand quality.

Another interesting finding (see Table 1) is that price does not appear among the best four cues to perceived quality for any of the five products. However, Table 2, which presents, for each product, the four attributes with the highest frequency of self-reported consideration in a brand purchase decision, reveals that price was a relatively important factor to consider when making a brand purchase in each of the five product categories. Thus, although these housewives felt that price was an important factor to be considered in making purchase decisions among brands, they did not feel price to be an especially accurate indicator of brand quality. Note that virtually the same statement can be made regarding the product attribute/quality cue of "brand name!"

TABLE 1

FOUR MOST INFLUENTIAL CUES TO PERCEIVED BRAND QUALITY^a
(based on mean self-reported rankings, n = 69)

Rank	Living Room Rug	Hair Dryer	Ground Coffee	Shampoo	Aspirin Tablets
1.	nature of the fiber (nylon, wool, etc.) 2.53 (62) ^b	guarantees and warranties 2.83 (54)	taste 1.60 (58)	special purposes (i.e., for dry hair, oily hair) 1.62 (55)	professional recommendation 1.65 (46)
2.	guarantees 3.43 (46) & physical appearance 3.43 (54)	manufacturer's reputation 3.07 (41)	brand name 2.82 (45)	past experience with shampoos 2.70 (44)	special ingredients (e.g., buffering, more pain relievers) 2.44 (32)
3.		model and style (hand model, bouffant style) 3.19 (47)	past experience with brand 2.87 (45)	special ingredients 2.77 (22)	past experience with aspirin 2.62 (37)
4.	maintenance ease 3.95 (57)	research findings (e.g., consumer reports) 3.42 (26)	type of grind (regular or drip) 3.24 (45)	hair dresser's recommendation 3.10 (20)	research findings 2.19 (23)

^a Ampersand (&) indicates a tie in the mean rankings.

^b Read as: Mean rank of "nature of the fiber" in terms of accurate indication of product quality was 2.53, based on the 62 people who considered "nature of the fiber" in their purchase decision.

TABLE 2

FOUR MOST IMPORTANT FACTORS TO CONSIDER
WHEN PURCHASING A BRAND IN FIVE PRODUCT CATEGORIES^a
(based on frequency of self-reported consideration, n = 69)

Rank	Living Room Rug	Hair Dryer	Ground Coffee	Shampoo	Aspirin
1.	nature of the fiber (nylon, wool, etc.) (62) ^b	guarantees and warranties (54)	taste (58)	special purposes (55)	professional recommendation (46)
2.	maintenance ease (57)	price (52)	price (49)	past experience with the shampoo (44)	past experience & brand name (37)
3.	price (56)	model and style (hand model, bouffant) (47)	brand name & past experience & type of grind (45)	brand name (40)	
4.	physical appearance (54)	brand name (44)		price (37)	price & special ingredients (32)

^a Ampersand (&) indicates a tie in the frequencies.

^b Read as: 62 (of 69) persons considered "nature of the fiber" in purchase decisions involving living room rugs.

Finally, although there was a lack of agreement across products in terms of both important quality cues and considered product attributes, some general correspondence between the two factors was evident within product categories. Moreover, this consistency extended beyond the most important attributes throughout the entire list. Table 3 presents Spearman rank order correlations between product attribute consideration in a purchase decision and attribute importance as a cue to quality, computed across all attributes for each of the five product categories. Clearly, the relationships are fairly strong, indicating a general tendency for consumers to use as cues to product quality those product attributes which are important considerations in purchase decisions, or vice versa. For the researcher, this relationship is a useful aid in identifying important quality cues, since salient cues are also likely to be product-related factors important in the general purchase decision-making process.

Table 3

Relationship Between Product Attribute Consideration in a Purchase Decision and Importance as a Cue to Product Quality

Product	Number of Attributes	Rank Order Correlation
Living Room Rug	15	.871
Hair Dryer	14	.764
Ground Coffee	12	.795
Shampoo	15	.653
Aspirin Tablets	13	.864

Conclusions and Discussion

Several tentative conclusions can be drawn from the present exploratory investigation. First, consumers may use several (e.g., 4-7) product attributes as evaluative criteria in brand-choice decisions rather than only 1 or 2. This finding suggests that future researchers should include several quality cues in combination in their experimental designs. Interestingly, prior research indicates that cues in combination may not have simple linear effects upon quality judgments, but rather may evidence non-additive effects through complex interactions between cues. For example, Olson (1972) determined that of the 11 quality perception studies in which cue interaction could occur, 9 investigations found such effects (e.g., Andrews & Valenzi, 1971; Enis & Stafford, 1969; Jacoby, Olson, & Haddock, 1971).

Second, given the finding that intrinsic cues were perceived to be somewhat more accurate indicators of product quality than were extrinsic cues, a tentative conclusion is that intrinsic cues have a more powerful effect upon judgments of quality than do extrinsic cues. This suggests that certain frequently studied extrinsic cues such as price, brand name, and store image are not expected to exert relatively strong effects on quality perceptions--

unless, of course, more potent cues are omitted from the study. Support for this conclusion is provided by work based upon and subsequent to the present study (Olson, 1972; Szybillo & Jacoby, 1972), as well as earlier experiments which included both intrinsic and extrinsic cues (Andrews & Valenzi, 1971; Jacoby, Olson, & Haddock, 1971). All of these investigations found that intrinsic cues (e.g., physical product differences such as taste) had stronger effects upon quality perceptions than did extrinsic cues (e.g., price, store image, brand name). Moreover, a reexamination of the experimental literature to date reveals that in most of the instances in which extrinsic cues had strong effects on quality perceptions, the more powerful intrinsic cues had been omitted from the study (e.g., Enis & Stafford, 1969; McConnell, 1968; Peterson, 1970). In addition, certain data indicate that the "true" state of quality cue influence may be even more complex than described above. Results obtained by Jacoby, Olson, and Haddock (1971) suggest that the effect of extrinsic cues on perceived quality may be evidenced primarily through interactions with intrinsic cues rather than through main effects. For these reasons, therefore, it is suggested that future attempts to experimentally assess the impact of specific extrinsic cues on perceived quality should also incorporate intrinsic cues in the design (especially the kinds of cues obtained through actual product usage--e.g., taste, appearance, fit, etc), if the results are to be externally valid and maximally meaningful.

A third conclusion, related to the second and partially supportive of it, is that, while price (an extrinsic cue) is an important factor to consider in a purchase decision, the price cue generally is not perceived to be as good an indicator of quality as are certain intrinsic cues (or other extrinsic cues). Support for this statement may be found in recent investigations which have found that price, when combined with other cues, has an insignificant impact on quality judgments (e.g., Jacoby, Olson, & Haddock, 1971; Szybillo & Jacoby, 1972; Valenzi & Andrews, 1971).

The final conclusion (actually a combination of several) is based upon both the present data and a critical review of the previous quality perception literature (see Olson, 1972). After a substantial amount of experimental effort (19 studies), we still know very little about the answers to the following questions: (a) Which product attributes, from among the many available, are chosen by the consumer as surrogate indicators of product quality? (b) Why are these particular cues chosen for use and others are not? (c) Once cues are chosen, how are they used (i.e., combined) to form a quality judgment? (d) Why are the cues combined in that manner--i.e., what factors determine cue usage and cue impact upon the quality judgment? It seems that our present lack of understanding is at least partially due to the atheoretical, exploratory basis upon which most investigations, including the present one, have been designed (see Cox, 1962, for a notable exception).

An obvious conclusion, therefore, is that the topic of quality judgment formation requires a theoretical model or conceptual framework which defines and describes the critical, determining factors in the quality perception process. At least three major elements are involved in this process--product, consumer-types, and quality cues (product-related attributes)--each of which may require its own model/framework. An attempt has been made to develop each of these models (Olson, 1972); however, the remainder of this paper will briefly present only that model concerned with quality cues.

A Model of Cue Choice and Usage

A basic assumption involved in the development of a model dealing with cue choice and usage is that, instead of simply examining whether specific cues affect a quality judgment for a certain product, a more fruitful research strategy is to begin at a more basic, abstract level and attempt to identify the factors which influence (a) whether or not a cue will be used in the judgment process, and (b) if so, the magnitude of cue impact on the final judgment. Such a conceptual model would provide a basis from which researchers could predict which cues affect quality perceptions and the relative strength of their effects.

Previously, few researchers have developed abstract constructs with which to describe and categorize quality cues in terms of dimensions relevant to cue usage and impact. Exceptions are Cox (1962) and Tull, Boring, and Gonsior (1964), although their work has not been followed up by other researchers. Two dimensions of the present framework were originally suggested by Cox and are modified and extended here.

The model presented below consists of two major dimensions (cue predictive value and cue confidence value) and, in addition, a dichotomous classificatory factor (cue intrinsicness/extrinsicness). Any potential quality cue can be described in terms of these three variables.

Predictive Value

Cue predictive value (PV) is defined as the degree to which an individual consumer associates a cue with product quality. That is, cue PV is the extent to which the consumer perceives or believes that the cue is related to or is indicative of product quality. To use a quantitative analogy, PV is similar to a probability coefficient which can take on values only between .00 (i.e., no probability that the cue is associated with product quality) and +.99 (i.e., near certainty that a cue is predictive of product quality).

It must be emphasized that PV is concerned only with the generic cue (i.e., price as a general cue, not specific levels of price) and the perceived predictive relationship between the generic cue and product quality. Although it is recognized that specific levels of a cue (e.g., high or low price) may indicate various degrees of quality (e.g., high or low quality), such considerations are not incorporated into the model at its present level of development. The functional relationship between levels of a cue and degrees of perceived quality is expected to be product specific and, therefore, is left to empirical determination.

Operationally, PV can be measured by having consumers rate the degree to which they believe a cue is indicative of product quality. For example, given the extrinsic cue of price:

How accurately do you think the PRICE of a brand of
beer indicates the OVERALL QUALITY of that brand of beer?

Not at all accurate	Slightly accurate	Moderately accurate	Very accurate	Extremely accurate
------------------------	----------------------	------------------------	------------------	-----------------------

○..... ○..... ○..... ○..... ○

Such ratings are obviously product specific and, therefore, PV ratings for the same cue may vary across products. Also, it should be emphasized that PV is measured at the level of the individual consumer and thus is not necessarily constant for a given cue and product across all consumers.

Confidence Value

The second major dimension of the conceptual cue framework is the confidence value (CV) of the cue, defined as the degree to which a consumer is confident in his ability to accurately perceive and judge that cue. Cue CV is the individual consumer's self-confidence in his ability to distinguish the cue and make accurate evaluations and judgments concerning it.

Confidence value can be operationalized by asking the consumer to rate his confidence in his judgment or evaluation of a given cue. For example, given the intrinsic cue of flavor:

How confident are you in your ability to perceive and evaluate differences in the beer samples in terms of their FLAVOR?

Not at all Confident	Slightly Confident	Moderately Confident	Very Confident	Extremely Confident
-------------------------	-----------------------	-------------------------	-------------------	------------------------

○..... ○ ○..... ○..... ○

Like PV, the CV of a cue is rated relative to a specific product by an individual consumer. For instance, a consumer may perceive a high CV for the cue of "taste" relative to wine, but a low CV for the "taste cue" relative to beer. In addition, the CV of a particular cue relative to a single product will vary across consumers as a function of differential learning experiences, cue knowledge, and cue familiarity. Furthermore, the average consumer is likely to perceive different CV's for the various cues related to a single product. Unless the consumer is an "expert" concerning that product, he is not likely to have high CV for all possible quality cues.

Awareness

An implicit aspect of the two-dimensional cue framework is awareness. Since the model has a distinctively cognitive orientation, it must be noted that the model applies only to those cues of which the consumer is consciously aware. That is, even though cues perhaps influence quality perception in a sub- or unconscious manner, the present model does not deal with such processes. Thus, in order to apply the model it is first necessary to identify which cues to perceived quality, of the many potential cues, an individual is aware of.

Intrinsic-Extrinsic

The intrinsic-extrinsic dimension (as described earlier) of the quality cue model is a dichotomous classificatory variable which has no direct effect on the process of cue utilization, but rather may be useful as a basis for determining the relative order in which individual cues from a set will enter the judgment process.

Combining PV and CV

The dimensions of PV and CV are both necessary to meaningfully describe and categorize quality cues in terms of cue utilization. It is assumed that these two dimensions alone (as mediated by the third classificatory variable--intrinsicness/extrinsicness) are sufficient to predict and account for quality cue choice and magnitude of cue impact upon quality perceptions. Based upon the ease with which one can imagine specific quality cues which are high on one dimension and low on the other, and vice versa, it seems reasonable, on a conceptual basis and pending empirical evidence, to treat PV and CV as orthogonal dimensions.

Hypotheses

Several working hypotheses have been derived from the conceptual model of cue utilization (see Olson, 1972), and others may be formulated based on future research. Following are three of the major hypotheses:

1. It is predicted that cue PV and CV in combination do not have an independent, additive effect on the probability of cue utilization or on the magnitude of cue effect. Rather, cue PV and CV are hypothesized to have an interactive effect on both aspects of the quality perception process. The general shape of this interaction is presented in Figure 1. Stated verbally it reads, only when a cue has both high PV and high CV does it tend to be used in the quality judgment process and have a strong effect on the final judgment.
2. It is also hypothesized that, for most products, consumers generally believe intrinsic cues are more accurate indicators of product quality (i.e., have high PV) than are extrinsic cues. Therefore, given that both intrinsic and extrinsic quality cues are available to the consumer and are perceived by him, it is predicted that intrinsic cues are used more often and, when used, have a greater effect upon quality perception than do extrinsic cues.
3. Finally, as a more specific extension of the previous hypothesis, it is predicted that the use of extrinsic (indirect) cues to product quality is related to the PV and CV of those intrinsic cues cognitively available to the consumer. Specifically, the tendency to use extrinsic cues in the formation of quality judgments is hypothesized to be an interactive function (see Figure 2) of the general PV and CV levels of the available intrinsic cues. Stated verbally it reads, extrinsic cues have a greater tendency to be used when available intrinsic cues have low PV, low CV, or both, and a lesser tendency to be used when intrinsic cues have high PV and CV.

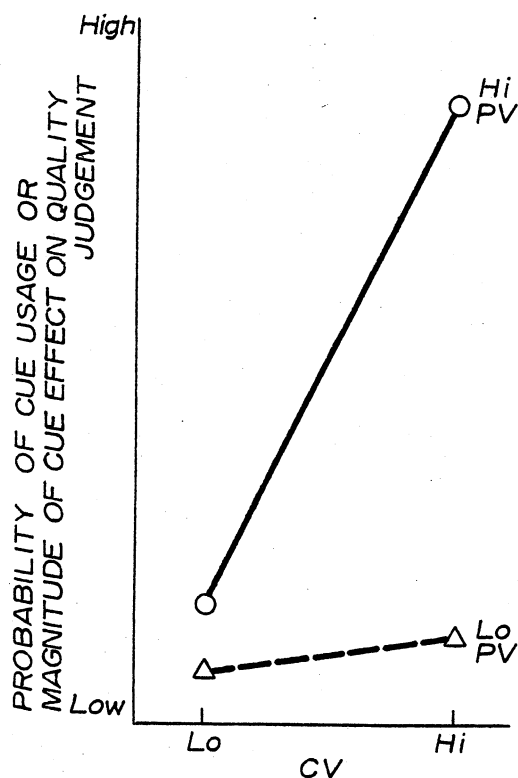


Fig. 1. Predicted PV X CV interaction effect on cue usage and cue impact.

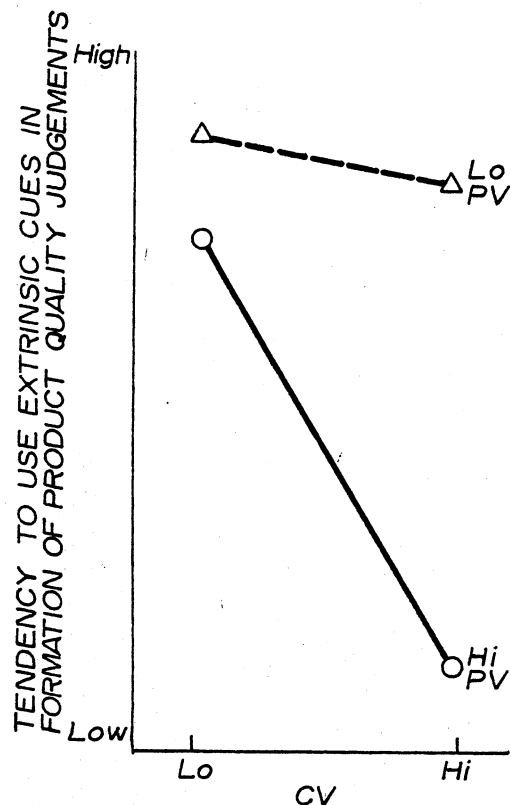


Fig. 2. Tendency to use extrinsic cues as a function of the PV and CV of available intrinsic cues.

Summary

This paper has presented an exploratory study intended to provide information regarding general relative importance of specific cues to product quality. Based upon the results of this study and a critical review of the perceived quality literature, a serious lack of theoretical and conceptual direction was noted in the quality research area. In an attempt to provide such a direction, we have briefly presented a conceptual model dealing with the process of quality cue utilization. Several working hypotheses derived from this framework were also briefly discussed. One can hope that researchers interested in developing an understanding of the process of quality perception will adopt a more abstract, conceptual approach (either the one presented here or one of their own construction) in future investigations of perceived product quality.

Footnotes

1. The authors gratefully acknowledge the assistance of Bruce Anderson in conducting this study.
2. Assistant Professor of Marketing, College of Business Administration.
3. Associate Professor of Psychology.

4. For example, Jacoby and Hollander (1972) obtained correlations of .775 (for orange juice), .763 (for bathroom tissue), .823 (for cooking oil), and .803 (for ketchup) for 97 housewives who responded (on 9-point semantic differential scales) to the following two questions: (a) "The brands of some products vary greatly in quality while the brands for other products hardly vary at all. About HOW LARGE do you feel the DIFFERENCES in QUALITY are BETWEEN BRANDS of bathroom tissue? Circle the number on the scale below that comes closest to your feeling on the matter." and (b) "When you do your shopping, how IMPORTANT is it for you to get the brand of (orange juice) that you USUALLY BUY? Circle the number on the scale below that comes closest to your feeling on the matter."

References

- Allison, R. I., & Uhl, K. P. Influence of beer brand identification on taste perception. Journal of Marketing Research, 1964, 1, 36-39.
- Andrews, I. R., & Valenzi, E. R. Combining price, brand name, and store cues to form an impression of product quality. Proceedings of 79th Annual Convention of the American Psychological Association, 1971, 6, 649-650.
- Cox, D. F. The measurement of information value: A study in consumer decision-making. In W. S. Decker (Ed.), Emerging concepts in marketing. Chicago: American Marketing Association, 1962, 413-421.
- Engel, J. F., Kollat, D. T., & Blackwell, R. D. Consumer behavior. New York: Holt, Rinehart, & Winston, 1968.
- Enis, B. M., & Stafford, J. E. Consumers perception of product quality as a function of various informational inputs. In P. McDonald (Ed.), Marketing involvement in society and the economy. Chicago: American Marketing Association, 1969, 340-344.
- Gardner, D. M. Is there a generalized price-quality relationship? Journal of Marketing Research, 1971, 8, 241-243.
- Jacoby, J. A model of multi-brand loyalty. Journal of Advertising Research, 1971, 11, (3), 25-31.
- Jacoby, J., & Hollander, S. W. Using dollar-metric scaling procedures to measure multi-brand loyalty. Unpublished manuscript, 1972.
- Jacoby, J., Olson, J. C., & Haddock, R. A. Price, brand name, and product composition characteristics as determinants of perceived quality. Journal of Applied Psychology, 1971, 55, 570-579.
- Laird, D. A. How the consumer estimates quality by subconscious sensory impressions. Journal of Applied Psychology, 1932, 16, 241-246.
- Leavitt, A. J. A note on some experimental findings about the meaning of price. Journal of Business, 1954, 27, 205-210.
- McConnell, J. D. The effect of pricing in an experimental setting. Journal of Applied Psychology, 1968, 52, 331-334.

- Olson, J. C. Product quality perception: A model of quality cue utilization and an empirical test. An unpublished doctoral dissertation, Purdue University, 1972.
- Peterson, R. A. The price-quality relationship: Experimental evidence. Journal of Marketing Research, 1970, 7, 525-528.
- Rao, V. R. Salience of price in the perception of product quality: A multi-dimensional measurement approach. In F. C. Allvine (Ed.), Marketing in motion/Relevance in marketing. Chicago: American Marketing Association, 1971, 571-577.
- Szybillo, G. J., & Jacoby, J. The relative effects of price, store image and composition differences on product evaluation. Purdue Papers in Consumer Psychology, Department of Psychological Sciences, Purdue University, 1972, Paper No. 123.
- Tull, D. S., Boring, R. A., & Gonsior, M. H. A note on the relationship of price and perceived quality. Journal of Business, 1964, 37, 186-191.
- Valenzi, E. R., & Andrews, I. R. Effect of price information on product quality ratings. Journal of Applied Psychology, 1971, 55, 87-91.
- White, I. S. The perception of value in products. In J. Newman, On knowing the consumer. New York: John Wiley, 1966.

THE RELATIVE EFFECTS OF PRICE, STORE IMAGE, AND
INTRINSIC PRODUCT DIFFERENCES ON PRODUCT QUALITY EVALUATION

George J. Szybillo and Jacob Jacoby

Department of Psychological Sciences
Purdue University

According to early economic theory, a consumer's choice was based on the assumption that he knew what he was buying. This was probably a valid assumption considering the limited range of consumer goods available at the time. With increasing technology, however, products have become more varied and complex. As a result of these changes, the average consumer today seldom has the time, knowledge, or the means to make a direct determination of product quality (cf. Cox, 1962; Olson, 1972; Shapiro, 1970).

Cox (1962) was one of the first investigators to develop a model of the consumer product evaluation process. He suggested that a product could be viewed as "an array of cues." The consumer's task in evaluating any given product was to use cues from this array for making evaluative judgments about that product. Cues were conceived of as relating to dimensions or attributes of the product, and included such information as price, brand name, store name, friends' and salesmen's opinions, and inter-brand composition differences.

Cox hypothesized that consumers tend to evaluate cues on two dimensions: predictive value and confidence value. Predictive value reflected the probability that the cue was associated with an attribute of the product. Confidence value reflected the certainty that the consumer felt about his ability to interpret and use that cue. For example, consider an audiophile and an unsophisticated housewife, both of whom are trying to form judgments about the quality of a tape deck. Both might share the belief that internal components represent a high predictive value. The audiophile, however, would probably have a higher degree of confidence than the housewife in judging whether the components were of good quality.

More recently, Olson (1972) has developed an extended conceptualization of the quality perception process. Quality perception is presented as a two-stage process in this model. Briefly, in Olson's Stage One the consumer perceives, distinguishes, and makes judgments of cues which he believes are related to product quality. Since it is difficult for the consumer to carry out objective product quality testing procedures, the average consumer judges product quality on the bases of cues which are "surrogate indicators." These indicators are variables associated with a product which are perceived by that consumer to be related to product quality. In Stage Two, the consumer combines his judgments of the cues into an overall judgment of product quality.

In addition to the two-stage description of the quality perception process, a conceptual framework for quality cues is provided which includes the concepts of: (a) intrinsic vs. extrinsic cues; and (b) cue awareness. Although other investigators (cf. Jacoby, Olson, & Haddock, 1971; Olson & Jacoby, 1972; and Valenzi & Andrews, 1971) have referred to the intrinsic-extrinsic dimension of quality cues, Olson (1972) has presented the most formalized approach to this dimension.

In this system, a cue to product quality is considered in terms of the degree to which it is intrinsically or extrinsically a part of the physical

product. A cue is considered intrinsic to the extent that, if changed, a resultant change would take place in the physical product itself. As examples, intrinsic cues to the quality of beer would be flavor and aroma, while extrinsic cues would include price and brand name.

Predictive value, confidence value, and cue awareness are applicable to both intrinsic and extrinsic cues in this system. In general, cues influence quality perception only if the consumer is aware of them (Olson, 1972).

A number of hypotheses have been derived from this conceptual framework. One of these states that, given both intrinsic and extrinsic cues, the consumer will depend more heavily on intrinsic than extrinsic cues in forming impressions of product quality. More specifically, extrinsic cues will have little effect on judgments of product quality when the predictive value and confidence value of intrinsic cues are high.

Data from two earlier studies tend to support this hypothesis. Valenzi and Andrews (1971) found that actual product composition differences in butter samples accounted for 13% of quality rating variance while price information accounted for only 4%. Jacoby et al. (1971) found that composition differences in beer samples accounted for 35% of the quality rating variance, considerably more than either price or brand name information.

The primary purpose of the present investigation was to provide a specific, a priori test of this hypothesis, using a product possessing different cues than were involved in the earlier studies. While the intrinsic cues employed by Jacoby et al. (1971) and Valenzi and Andrews (1971) were primarily olfactory and gustatory in nature, the intrinsic cues for the product used in this study (ladies nylon hosiery) were primarily tactile.

Extrinsic cues utilized in the present study were price and store image information. Only two studies have investigated the effects of price and store image in combination as cues to product quality (Andrews & Valenzi, 1971; Stafford & Enis, 1969). Both studies were concerned with the specific effects of various levels of the two cues. The present study investigated the relative potency of these cues rather than focusing on the effects of various levels of these cues. Moreover, neither of these studies permitted actual product differences to vary as a third factor in the experimental design. Recent research (Jacoby et al., 1971) suggests that such procedures may over-simplify the quality perception process. Another consideration is that subjects in both these studies were exposed to all experimental conditions in their respective studies. This suggests that confounding demand characteristics may have unknowingly been introduced. In contrast, subjects in the present study were exposed to only a single experimental condition, i.e., there was independence of subjects across cells.

Finally, Shapiro (1970) has recently presented a model of quality perception for products without brand names. The model is based on an analysis of the interrelationships of perceived quality, perceived likelihood of purchase, and perceived worth (i.e., value for the money) for sweaters, chairs, and hosiery. Shapiro found that, for these products, perceived worth had a stronger relationship to perceived likelihood of purchase than did perceived quality. Inasmuch as unbranded product samples were used in the current investigation, the relationships between perceived quality, likelihood of purchase, and perceived worth were also examined. The specific hypotheses of the study were:

Hypothesis 1. Intrinsic cues (i.e., cues which are directly due to composition differences between brands) would exert a greater effect on quality perception than would extrinsic quality cues (e.g., price or store image).

Hypothesis 2. Perceived worth, defined as value for the money, would have a stronger relationship with perceived likelihood of purchase than would perceived quality.

Method

Subjects

The subjects (Ss) were 90 female undergraduates at Purdue University enrolled in Introductory Psychology during the fall 1971 session. One restriction was placed on subject participation: subjects were required to have lived in the Lafayette area for a minimum of six months. The purpose of this requirement was to insure that the independent variable of store image possessed meaning for the Ss. (N.B. An additional 23 and 14 Ss, respectively, were used to pre-test the instructions and experimental manipulations.)

Design

A 2 X 3 X 3 factorial, consisting of two levels of price (i.e., present vs. absent), three levels of store image (i.e., high quality store image, low quality store image, and no information provided as to store image), and three levels of product samples (i.e., three physically different samples of panty hose), was employed. Repeated measures were taken across the third factor (Winer, 1971). Fifteen Ss were used in each of the six experimental conditions.

Independent Variables

Price (P) information was manipulated by either presenting or withholding the actual prices of the three product samples. The prices employed (\$0.98, \$1.67, and \$2.49) were based on the results of one of the pre-tests as well as comparative shopping.

Store image (SI) information was manipulated by either presenting or withholding the names of stores purportedly selling the three product samples. Two store names were selected on the basis of the pre-test to represent high and low quality store images. The difference between the store with the highest mean quality rating ($\bar{X}_H = 7.95$, on a 9-point scale) and the store with the lowest mean quality rating ($\bar{X}_L = 2.56$) was highly significant ($t = 16.58$, $df = 22$, $p < .001$).

Each product sample (PS) of panty hose was identified by either I, J, or R. This decision was based on a pre-test which showed no differences in preference between these three letters. The letters were randomly assigned to samples for each experimental condition. For purposes of a simplified exposition, the letters I, J, and R will be used to designate the low, medium, and high quality samples, respectively, throughout the remainder of this paper.

The three samples of panty hose were pre-tested on the attributes of knit, texture, workmanship, cut, and overall quality. Results of the analysis of variance of overall perceived quality indicated a significant main effect for samples ($F = 24.97$, $df = 2, 26$, $p < .0001$), and Newman-Keuls analysis indicated

that all differences between pairs of means for the three samples ($\bar{X}_I = 34.28$, $\bar{X}_J = 54.64$, $\bar{X}_R = 84.50$) were significant beyond the .01 level. These differences were in the anticipated direction.

Dependent Variables

The major dependent variable for this study was overall perceived quality. The minor dependent variables were perceived worth and perceived likelihood of purchase. All were measured using the 100-point scale procedure described in Jacoby et al. (1971).

Procedure

Each of the six cells was run separately. The subjects were seated at a long table and separated by table-top partitions to cut off between-subject communication. A female source introduced the experiment as a psychological study designed to develop reliable instruments for measuring hosiery perceptions.

Subjects were then given the questionnaire booklet, and the female experimenter demonstrated how to use the rating scales. Subjects were told they would be evaluating three samples of panty hose available in the local area. Hosiery samples were presented in wrapped gift boxes. Price information and store image information were given, as appropriate, by labels attached to the experimental samples and communicated verbally before subjects began evaluating the hose. Samples were checked after each experimental session for damage, replaced if necessary, and new labels assigned.

Results

Manipulation Check

The manipulation check for store image indicated a significant main effect for "quality of store merchandise" ($F = 53.02$, $df = 2, 87$, $p < .001$). A Newman-Keuls analysis indicated that all differences between pairs of means for the three store image conditions ($\bar{X}_L = 3.80$, $\bar{X}_O = 5.73$, $\bar{X}_H = 7.93$) were significant beyond the .01 level. These differences were in the anticipated direction, with the store names selected on the basis of the pre-test to represent low store image and high store image receiving the lowest and highest mean ratings, respectively.

Model for Data Analysis of Dependent Measure

The model for a repeated measure analysis of variance is appropriate if the following assumptions are satisfied: (a) homogeneity of the variance-covariance matrices, and (b) the symmetry of the pooled variance-covariance matrix (Winer, 1971). For the dependent measure, tests for homogeneity of variance for variation due to subjects within groups and due to the interaction between samples and subjects within groups were computed. Complete assumptions about the variance-covariance matrices were not tested. Consequently, the analysis used the Greenhouse and Geisser (1959) procedure for all tests involving the samples factor. This procedure is negatively biased in that it uses conservative degrees of freedom, leading to errors in the direction of not rejecting false null hypotheses.

Data Analysis of Major Dependent Variable

The means and standard deviations of overall perceived quality ratings for each experimental cell are presented in Table 1. Results of the analysis of variance based on these data are summarized in Table 2. Tests for homogeneity of variance yielded insignificant results for variation due to subjects within groups ($F_{\max .95} = 4.24, df = 6, 14$), and due to interaction between samples and subjects within groups ($F_{\max .95} = 2.95, df = 6, 28$).

Table 1

Means and Standard Deviations of Overall Perceived Quality Ratings

	Price Present			Price Absent		
	Sample I	Sample J	Sample R	Sample I	Sample J	Sample R
Low Image:						
\bar{X}	24.80	36.46	85.13	19.80	49.86	82.46
SD	20.35	2.64	11.33	13.34	12.87	14.21
No Image:						
\bar{X}	26.46	54.73	82.73	24.73	51.33	86.86
SD	15.25	13.04	15.60	12.43	14.45	7.42
High Image:						
\bar{X}	30.00	57.33	87.06	34.40	56.86	86.66
SD	18.51	15.33	9.98	23.74	17.36	13.37

Note:-- n = 15 per cell.

Table 2

Analysis of Variance of Overall Perceived Quality Ratings

Source	df	MS	F
Between Subjects:			
Price (P)	1	56.948	<1
Store Image (SI)	2	1810.737	5.468**
P X SI	2	29.470	<1
Subj. W. Grps.	84	331.134	
Within Subjects:			
Samples (PS)	2	77582.893	394.859***
P X PS	2	93.359	<1
SI X PS	4	282.537	1.438
P X SI X PS	4	417.581	2.125
S X Subj. W. Grps.	168	196.482	

** p < .01

*** p < .001

A strong Product Samples main effect was obtained ($F = 394.85$, $df = 1, 84$, $p < .0001$), and a Newman-Keuls analysis indicated that all differences between the pairs of means for the three samples ($\bar{X}_I = 26.76$, $\bar{X}_J = 51.10$, $\bar{X}_R = 85.15$) were significant beyond the .01 level. The analysis also indicated a significant Store Image main effect ($F = 5.46$, $df = 2, 84$, $p < .01$). A Newman-Keuls analysis indicated that only the difference between the means for the low store image ($\bar{X}_L = 49.75$) and high store image ($\bar{X}_H = 58.72$) conditions was significant.

Intrinsic quality cues were anticipated to have a greater effect on quality perception than extrinsic cues. To test this hypothesis, ω^2 values were computed for the effects of price, store image, and product samples (Winer, 1971). In terms of these indices, 73% of the total variance was due to the main effect of samples, and approximately 1% was due to the main effect of store image. Price information was relatively ineffective in explaining any percentage of the total variance. This finding confirms Hypothesis 1.

The price-present, store-image absent condition of this study was analogous to Shapiro's (1970) experimental setting. The correlations between perceived quality and perceived likelihood of purchase for samples I, J, and R were .07, -.03, and .49, respectively. The correlations between perceived worth and perceived likelihood of purchase for samples I, J, and R were .72, .56, and .64, respectively. Tests were made for differences between dependent correlations (Bruning & Kintz, 1968). Perceived likelihood of purchase was more closely related to perceived worth than perceived quality for samples I and J ($t_I = 2.68$, $t_J = 2.60$, $df = 12$, $p < .05$), but not for sample R ($t_R = .75$). These results partially confirm the second hypothesis.

Discussion

Consistent with earlier research (Valenzi & Andrews, 1971; Jacoby et al., 1971), intrinsic quality cues were found to have a greater effect upon quality perception than did extrinsic cues. More specifically, composition differences in hosiery samples had a greater effect (accounting for 73% of the quality rating variance) than did either price or store image information. In fact, contrasted to previous studies, the variance accounted for by product composition differences in this study is relatively large. This may be an effect of moving from products (i.e., butter and beer) where between-brand differences, although discriminable, have a relatively small range, to a product (i.e., hosiery) where product differences can have a relatively large range.

Comparing the relative strength of price and store image, the results indicate that while the effect of price was not significant, awareness of store image did influence quality perception. Apparently, knowing the samples came from a store which carries high quality merchandise increases quality ratings relative to knowing that the samples came from a store which carries low quality merchandise. Although the store image effect was statistically significant, its practical importance is mitigated by the amount of variance it explained. Once again, the price main effect was insignificant when price was embedded in a multi-cue framework.

Shapiro's (1970) hypothesis that perceived worth would be more strongly related to perceived likelihood of purchase than would perceived quality received partial support. Given continued support, several substantial implications for practitioners would ensue.

Further research on the quality perception process is needed. Various writers have posited hypotheses regarding consumer characteristics and their

relationship to quality perception (e.g., Jacoby et al., 1971; Olson, 1972; Shapiro, 1970). Few, if any, of these hypotheses have been empirically examined. The conclusions of the present study and prior studies may require modification as more information is obtained regarding how such variables as frequency of usage, frequency of purchase, general self-confidence, perceived quality differences between brands, and importance of the product category affect the quality perception process.

References

- Andrews, I. R., & Valenzi, E. R. Combining Price, Brand Name, and Store Cues to Form an Impression of Product Quality. 79th Annual APA Proceedings, 1971, 6, 649-650.
- Bruning, J. L., & Kintz, B. L. Computational Handbook of Statistics. Glenview, Illinois: Scott, Foresman, and Company, 1968.
- Cox, D. F. The Measurement of Information Value: A Study in Consumer Decision-Making. In W. S. Decker (Ed.), Emerging Concepts in Marketing. Chicago: American Marketing Association, 1962, 413-421.
- Greenhouse, S. W., & Geisser, S. On Methods in the Analysis of Profile Data. Psychometrika, 1959, 24, 95-112.
- Jacoby, J., Olson, J. C., & Haddock, R. A. Price, Brand Name, and Product Composition Characteristics as Determinants of Perceived Quality. Journal of Applied Psychology, 1971, 55, 570-579.
- Olson, J. C. Product Quality Perception: A Model of Quality Cue Utilization and an Empirical Test. Unpublished doctoral dissertation, Purdue University, 1972.
- Olson, J. C., & Jacoby, J. Cue Utilization in the Quality Perception Process. Purdue Papers in Consumer Psychology, Paper No. 122, 1972.
- Shapiro, B. P. The Effect of Price on Purchase Behavior. In D. L. Sparks (Ed.), AMA Fall Education Conference. Chicago: American Marketing Association, 1970. P. 42.
- Stafford, J. E., & Enis, B. M. The Price-Quality Relationship: An Extension. Journal of Marketing Research, 1969, 6, 456-458.
- Valenzi, E. R., & Andrews, I. R. Effects of Price Information on Product Quality Ratings. Journal of Applied Psychology, 1971, 37, 106-191.
- Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill Book Company, 1971.

THE NEED FOR PROBLEM DEFINITION AND
RESEARCH EVALUATION OF PROPOSED SOLUTIONS FOR MAKING
PUBLIC POLICY DECISIONS RELATING TO MARKETING AND CONSUMER BEHAVIOR

Raymond C. Stokes
Director
Consumer Research Institute

Introduction

As a way of leading up to the ideas which I wish to discuss with you today, let me first briefly mention a phenomenon which is common knowledge to all of us. During recent years we have witnessed a new wave of consumerism represented by a fresh crop of consumer advocates who are more plentiful, devout and articulate than their predecessors. Their influence has already extended over a longer period than previous generations and their impact has been wider and deeper. While we may have witnessed the crest of the current wave of consumer activism, it has by no means run its course. It is impossible to predict how long the current movement will continue, but most would agree that it will persist for several years and some believe it will be permanent.

Responses to the Pressures of Consumerism

The forces of consumerism have already had a substantial impact on the marketing practices of business and the activities of governmental regulatory and legislative bodies. The regulatory activities of government agencies are being conducted with increased vigor especially at the Federal Trade Commission. Legislative bodies at the Federal, State and local levels have been busy formulating bills designed to protect the consumer. Many have been passed and hundreds are pending. In addition to increased governmental activity, there have been widespread voluntary efforts of business, such as unit pricing, open dating, nutrient labeling and policing of abuses in advertising.

Another unique and constructive response to consumerism has been the formation of the Consumer Research Institute whose purpose is to sponsor or conduct research in any area of marketing practice which may be the subject of consumer concern, for the purpose of helping in the shaping of public policy. CRI was formed about four and one-half years ago by grocery manufacturers, advertising agencies, media organizations, trade associations representing various segments of the grocery industry and several management consulting and market research firms. In so far as possible, CRI's research has been coordinated with appropriate governmental agencies and conducted in a time frame which would generate information useful to policy decision making. We have demonstrated that consumerism issues can be researched, that government officials will accept such research and that these facts can and will be used in making public policy decisions.

The Failures of Consumerism

While the dedication and concern of the consumer advocate must be admired, the increased vigor of the government commendable, the voluntary activities of business encouraging and the research on public policy issues initiated by CRI a desirable beginning, I believe an honest and objective evaluation of consumerism would conclude that it has so far yielded minor benefits to consumers at cost and therefore must be considered a failure. The consumer advocate and business apparently agree on one point: recent government regulation of the marketing function has been a substantial failure. Yet the consumer advocate proposes additional governmental control with the full knowledge that previous regulation has been less than satisfactory.

Consumer Problem Definitions and Proposed Solutions

It appears to me that one of the primary, if not the major, reason for the low yield and high cost of current consumerism activity is related to our lack of understanding of the problems and the adequacy, feasibility and cost of proposed solutions. Those knowledgeable of the marketing system and consumer behavior have not accepted the responsibility of defining problems and developing alternative solutions. Therefore, by default the consumer advocate, who is naive and unsophisticated in the complexities of marketing and consumer behavior, does his best to define our problems for us and to offer solutions he considers appropriate. It appears to me that the consumer advocate is pointing out symptoms rather than the fundamental underlying problems. Proposed solutions are superficial in that they only address the symptoms. Due to our zeal to be socially responsible citizens, we have been too quick to accept the advocate's definitions of problems and to implement his proposed solutions. Resources (especially research funds) are too limited and the problems too urgent for this to continue.

Examples of Inadequate Problem Definitions and Proposed Solutions

There are many examples of inadequate problem definitions and the proposal of superficial solutions, but I will restrict myself to a few with which the Consumer Research Institute has been most concerned.

Unit Pricing

Because of brand, product and package size proliferation along with fractional ounce packaging, it is difficult for consumers to compare the price of competing grocery products. The proposed solution is the posting of unit prices at the point of purchase. The assumptions are that consumers can use this information, will use it, and the cost of furnishing it will be less than the benefits which accrue. While the results are not all in, it is clear that the system will not reach those who need it most--the poor and undereducated. It also appears that actual usage is minimal where it has been installed and costs are high in comparison with the benefits actually obtained.

Open Dating

The consumer advocate recently became aware that it was good manufacturing practice to imprint or deboss a code on each individual food

package indicating the day, shift, manufacturing plant, packaging line and perhaps other information, of significance for quality control purposes. While the advocate had no interest in decoding most of this information, he wondered why the date of manufacture or packing couldn't be communicated in plain english so we could all know the age of the food. He was quick to believe that there was an excessive amount of stale or bad food offered for sale in grocery outlets in this country and that the printing of a date on the food/package would automatically solve the problem. No objective evidence was offered of the magnitude of the problem of stale food and certainly no evidence was offered that open dating would solve the problem. The Consumer Research Institute has been engaged in research activities in this area and finds that there is a favorable psychological reaction to the installation of open dating, but there is no evidence of improvement in the freshness of the food offered for sale. I view open dating to be more of a placebo than a panacea. Benefits which are promised but not delivered could contribute to additional consumer mistrust of the marketing system and ultimately result in increased consumer dissatisfaction.

Nutrient Labeling

Studies have shown that certain segments of our society have nutritional deficiencies, though we have the most abundant, wholesome food supply in the history of mankind. The White House Conference on Food, Nutrition and Health recommended that nutritional information be printed on food packages. The Food and Drug Administration and the leaders of the food industry enthusiastically agreed and the Consumer Research Institute has been engaged in a cooperative research project with the FDA to help determine the format for communicating this information which would be best understood and most useful to consumers. I know of no one who disagrees with this move. However, I have never and I do not now believe that this is a realistic solution to the problem of nutritional deficiency. The information is being communicated at a time, at a place, in a form which insures that it will have minimum usefulness to the consumer. The shopper in the supermarket does not have the time to run up and down the aisles reading and comparing nutritional statements and planning menus which will be nutritionally balanced. Clearly, we need a much better understanding of the problem, particularly from a consumer behavior point of view before we can begin to think intelligently about solutions. It does appear to me, however, that it might help to have additional nutritional information available to the consumer in the home, perhaps in the form of a nutritional handbook which would list nutritional values, not only by generic types of food but by brand name as well.

Restrictions of Advertising to Children

Recently, a group of concerned mothers in Boston organized a group called Action for Children's Television (ACT) which petitioned the Federal Communications Commission to ban all advertising directed to children on the assumption that it was doing various kinds of harm. The Consumer Research Institute asked the Associates for Research in Behavior to examine and conceptualize these charges, survey the literature to learn what is known that would speak to the charges and to suggest needed research. Their analysis indicated that the charges leveled against the advertising industry are unsubstantiated by any known factual information. Since this is such a serious charge, the Consumer Research Institute has recently sponsored a seminar under the chairmanship of Seymour Banks of the Leo Burnett

Advertising Agency which brought together a group of social scientists to define the problems and attempt to outline a long-range research program on "How Children Learn About Being a Consumer." This group of child development psychologists, social psychologists, sociologists, communications and marketing experts spent a day and a half in discussions, a transcript of which will be made available to all who are interested in conducting research on the consumer socialization of the child.

Recommendations in Regard to Defining Problems and Developing and Evaluating Alternative Solutions

I hope you will agree that so far we have not taken the time and trouble to dig deeply enough to define the basic, fundamental, underlying problems in marketing and society which are contributing to the new wave of consumerism. If this is true, it seems clear that we must begin to devote a larger portion of our time and money to problem definition. I suggest that the following approaches may be productive:

(A) Hold seminars similar to the Seymour Banks conference on the consumer socialization of the child mentioned above, in which the interested and knowledgeable experts in the country would define the real problems, suggest possible solutions and outline research programs. For example, two seminars seem to be needed on the difficult problems associated with the public policy aspects of advertising. One on the "copy" or psychological aspects of advertising, particularly the phenomenon of deception. I suggest that such a meeting, which would be attended by representatives of the FTC, advertisers, advertising agencies and academic, be held shortly after Professor John Howard finishes his summary of the recent FTC hearings on advertising. The second seminar would consider the economic aspects of advertising. It may be appropriate for organizations like the American Enterprise Institute for Public Policy Research, the Brookings Institute or a university economic department to sponsor a seminar of this type in which those economists antagonistic to advertising would be brought together with those who are pro-advertising to agree upon the kinds of data which will be required and methodologies which would help resolve their differences.

(B) Direct specific research activity toward the conceptualization and definition of fundamental problems. CRI is sponsoring one such study at present with Professor Robert Perloff and his associates at the University of Pittsburgh. This group will conceptualize the various dimensions or aspects of the phenomenon of deception, develop operational definitions and suggest ways in which deception may be reliably measured. A second phase of the study will be the development and application of instruments to measure deception. It would be desirable to objectively measure deceptiveness of advertisements rather than depending upon legal or expert judgment. This would also allow a more objective measure of the effectiveness of corrective and counter advertising.

(C) Another approach which I feel should be helpful is intensive, in-depth study of the environment in which the problems are believed to exist. For example, Professors Jagdish Sheth, Peter Wright and their associates at the University of Illinois are preparing a proposal for CRI to study the modern american supermarket. They would first develop a topology of the informational display which consists of around 8,000 items of each of which has many bits of information in addition to other promotional types of material in this environment. It may well be that the modern american supermarket is already the most complex informational display

experienced by large numbers of people on this planet. The second phase of this investigation would be designed to develop a better understanding of the information processing and decision making task faced by the shopper.

In parallel with this proposed in-store study, CRI is sponsoring a series of laboratory projects with Professor Jacob Jacoby and his students at Purdue University to determine the amount of information actually used by consumers in making grocery purchase decisions and the hierarchy of importance or salience of this information. These studies should allow us to more adequately define problems with the marketing of groceries and develop alternate solutions for evaluation.

(D) Governmental and consumer advocate contact with consumers comes primarily from complaint letters. Too frequently such letters are from a highly vocal minority who may not be representative of American consumers. I urge that we commence conducting large scale surveys on a periodic basis to develop measures of consumer satisfaction, consumer expectations and perceived failings of our marketing system. Such information would not only allow better diagnosis of problems and the development and implementation of voluntary or Governmental solutions, it would also yield data for evaluating the effectiveness of corrective action.

Some of us have been accused at times of using research as a delaying tactic. I hope those who make this charge will allow marketing and consumer behavior scientists to spend the necessary time to do sound research. I hope governmental and legislative bodies will not rush into further regulation and control until the necessary research is done and facts are available upon which sound public policy decisions can be based. This, of course, will require a considerable amount of money which is not now available. The Consumer Research Institute has made a start, but we will be unable to support the level of research which will be required in the future. Let us hope that others, especially foundations and governmental agencies, will recognize this problem and support research activity on the public policy aspects of marketing and consumer behavior.

A CONCEPT OF CONSUMER POLICY

Hans B. Thorelli
Indiana University

"Consumer policy" is a simple translation from konsumentpolitik, a term originating in highly consumer-conscious Scandinavia. On neither side of the pool has any real attempt been made to analyze what it is all about. Our objective here is to take a first step in this direction. Operationally, we may define consumer policy as measures taken to implement consumer interests. The delivery system of consumer policy comprises consumer education, consumer information and consumer protection.

Let us see where these initial statements may lead us.

The Consumer Interest and the Public Interest. The Freedom to Consume

A helpful start may be to relate the consumer interest to the public interest.¹ It may sound trite to say that the public interest manifests itself in what the public and/or its representatives do by way of passing (or not passing) laws, rules and regulations and taking (or not taking) decisions and actions in matters of political concern. For our purposes this is a sufficient definition, however, and volumes of political philosophy indicate that it is actually difficult to give the concept much more meaning and still retain a broad acceptance. The consumer interest, on the other hand, is perhaps most readily identified with established, or at least articulated, consumer rights. Paradoxically, the most important consumer right is hardly ever mentioned. This is the "freedom to consume," a right which is greater than and antecedent to the four consumer rights enunciated a decade ago and since quoted thousands of times around the world.

We need to do no more than contrast the freedom to consume with the public interest as manifested in laws and official action to realize that the consumer interest and the public interest are not identical.² This is easily seen in a country like the USSR where consumption has been deliberately restricted for decades in favor of military and industrial development, and in rationing systems artificially restraining private consumption in Western countries in World War II.³ It is no use protesting that "we are all consumers." Every man wears several hats and plays several roles, and by definition no single role (except perhaps that of the voter) can adequately express the public interest. Indeed, conceptually most consumers are also producers--and few have had the temerity to suggest that the producer interest is equal to the public interest.

The consumer interest, like any "special" interest, is a subset of the public interest. What is good for the consumer may not always be good for the country. Nevertheless, as it concerns everybody, the consumer interest is clearly a subset of grand importance. And to those of us who believe that inextricable links exist between political democracy and free markets the consumer interest is a matter of paramount concern.

It must be admitted, however, that to define "the consumer interest" in specific situations tends to be a task fraught with ambiguity. The fractionalization of the consumer interest begins right in the marketplace. Preferences with regard to style, quality, and price, as well as the total bundle of

products desired, vary greatly among consumers of different income, age and educational groups and frequently even within these groups. The open market is the equilibrating mechanism among all these consumer "interests." Beyond the marketplace, some consumers are ecologically oriented in the process of consumption and the disposal of refuse while others are not, etc. Indeed, it is precisely because our role as consumers is such an all-pervasive aspect of life that the consumer interest is so differentiated. No doubt this is why it has proven impractical to bring about a unified consumer movement of major proportion in most countries. Of course, this does not mean that consumerism is doomed to be a passing fad rather than an ongoing social and political force. On the contrary, it may well grow in importance as the phenomena in which consumerism originates are likely to manifest themselves even more dramatically in the future. That is, the increasing complexity of the marketplace due to the proliferation, rapidity of change and technical intricacy of market offerings occurring in parallel with a seemingly forever widening functional distance between producer and consumer. That is, too, a revolution in consumer aspirations which--somewhat paradoxically--pre-industrial and post-industrial society seem to have in common.⁴

Consumer Policy-Makers

The consumer interest is in part a subset of the public interest. Thus, consumer policy is in part a subset of public policy. Government is clearly an important maker of consumer policy. It should be equally obvious it is not the only one: consumer policy is also in part a subset of private policy. Individual citizens have their own "consumer policies" as they pursue their interests in the marketplace. Of greater interest here, however, are policies affecting collective consumer interests. Makers of consumer policy in this broader sense include consumer organizations, other citizen groups, business, educational institutions and the mass media.

The appropriate roles of various policy-makers is a highly controversial topic to be given a once-over-lightly in a later section.

Typology of Consumer Policy: Information, Education and Protection

In a general sense, consumer information comprises all data about individual markets and offerings. Consumer information originates in personal experience, in commercial communications and in independent sources, notably comparative testing, informative labeling and quality certification. It is oriented to specific buying decisions. By contrast, consumer education may be thought of as "consumer civics": consumer education provides the knowledge foundation necessary to develop citizens into intelligent consumers, or at least to make their self-development into intelligent consumers possible. Thus consumer education extends all the way from conveying an understanding of how the market economy operates, of the consumer decision-making process and of consumer rights and responsibilities to such pragmatic matters as the properties of different textile fibres and dietary concerns. Consumer protection are measures (typically taken by others than the individual consumer) to safeguard consumer rights. Consumer protection ranges from competition (antitrust) policy to maintain open markets via the control of deceptive practices and the handling of consumer complaints to standards and other rules and regulations to maintain consumer health and safety.

Clearly, the distinctions between consumer information, education and protection are not hard and fast. For instance, the same information may have multiple uses. Americans will find the phrase "The Surgeon General has deter-

mined that cigarette smoking is dangerous to your health" on each and every cigarette package. That the message is intended for consumer protection is self-evident. But it could also be viewed as educational. And it could also be regarded as consumer information about any single brand you might happen to select. As regards the border area between consumer education and information we may say that the more "generic" the data are in terms of product or consumer characteristics the more likely it is that consumer education is the appropriate term, and that the more specifically they are related to individual offerings (brands) or to the needs of individual consumers, the more appropriate it is to speak of consumer information.

Consumer Rights and Consumer Policy: A Scenario

We would suggest that the nature and scope of consumer policy can be understood most readily by viewing it in the context of consumer rights and responsibilities. The graphic representation in Figure 1 will serve as a point of reference. The matrix arrays aspects of policy in the left-hand column and rights and responsibilities in the right-hand columns. If the book-page were more suited to three-dimensional diagrams we would have included consumer policy-makers as a logical third dimension.

In our typology of consumer policy we noted that the distinctions between education, information and protection are not hard and fast. The classification of consumer rights, seemingly first made in an address by J. F. Kennedy, is pretty traditional a decade later. It leaves out the "freedom to consume" (or not to consume) as being antecedent to, and more general and basic than those detailed in the figure.

The matrix tries to make two vital points, namely

that the enforcement of literally every consumer right logically depends on all three types of consumer policy, that is, education, information and protection, and

that for every consumer right there is a corresponding consumer responsibility.

The lesson here is simple and crucial: no matter how aggressively we may use consumer policy, it will not in itself suffice to enforce consumer rights. In the end consumer rights will exist only if at least some individual consumers really exercise some of their rights and responsibilities at least some of the time.

Positive and negative rights and responsibilities. The last statement suggests a few remarks on positive and negative rights and responsibilities. Rights and responsibilities are inherently two-faced. To the positive right to choose between viable alternatives (including having access to suppliers of the brands in question) corresponds the negative right not to have choice imposed (such as by some central planning authority, or by overly well-meaning consumer "protectionists"). To the right to be informed corresponds the right not to be deceived.⁵ To the right to be heard corresponds the right to privacy.⁶ To the right to safety corresponds the right to take some safety risks--such as smoking cigarettes or driving or swimming without a safety belt--at least as long as we are aware of the risks and are respecting the rights of other consumers.

Figure 1 Consumer Policy and Consumer Rights and Responsibilities

CONSUMER POLICY	CONSUMER RIGHTS			
	1. CHOOSE FREELY	2. BE INFORMED	3. BE HEARD	4. BE SAFE
A. EDUCATION	decision-making, budgeting; nature of market economy, rights and responsibilities	generic product and materials data, information sources	how to assert consumer rights	importance of health and safety, user manuals and training
B. INFORMATION	buying criteria buying advice	models and brands data, independent consumer info programs	market research, two-way market dialogue	safety certification, care and maintenance data
C. PROTECTION	maintain open markets, antitrust; stop hi-pressure and deceptive tactics	truly informative advertising, product claims substantiation	complaints handling machinery	minimize health and accident risks
	CHOOSE WISELY	KEEP INFORMED	SOUND OFF	SAFETY FIRST
	CONSUMER RESPONSIBILITIES			

A third dimension of the matrix would show the makers of consumer policy. These policy-makers include consumer organizations, other citizen groups, business, government, educational institutions and the mass media.

Consumer responsibilities present us with an analogous situation. The positive imperative to choose wisely is negated by our freedom to choose what we know is "wrong" for us, or to spend our money foolishly.⁷ To the positive duty to keep informed corresponds the "negative" phenomenon of impulse buying, or of taking a deliberate chance at an auction. At least as yet no voice has been raised to prohibit these types of behavior. To the duty of sounding off corresponds the urge to keep quiet, to avoid the unpleasantness and the waste of time involved in asserting our rights. And instead of battling for safety first at all times some of us at least occasionally will put performance⁸ or, indeed, even fun⁹ above safety. Or we may use a product for another purpose than that for which it was made.¹⁰ Or we may simply neglect maintenance, which in and of itself may be enough to bring safety hazards. The point of all this is that we cannot expect most consumers to be vigilantes of the marketplace most of the time.

It may be observed that most of us have quite positive--often adamant--feelings about what here is loosely termed the negative rights and responsibilities of consumers. Indeed, they all seem to represent widely embraced Western ideals. For the enforcement of positive consumer responsibilities society relies almost exclusively on voluntarism among consumers themselves. On the other hand, the positive consumer rights, while perhaps self-evident in theory, seemingly need to be constantly reasserted to stay alive. Even though the reassertion of these consumer rights is logically the prime responsibility of consumers, governments and other groups are having to add their weight. The market economy may create the potential for the realization of positive consumer rights but in a complex society there is no automaticity about their actual realization. This implies cost and effort on the part of individual consumers and consumer policy-making organizations in the struggle to enforce the rights.¹¹

Consumer policy measures. The cells in Figure 1 give some examples of consumer policy measures. They serve illustrative purposes only. There is nothing sacred about the arrangement; in several instances it is a matter of taste and emphasis rather than principle. This is almost inevitable, given the overlaps between education, information and protection as well as between the several consumer rights. The right half of the matrix would seem fairly self-explanatory. We shall comment on some of the rights and responsibilities in the left half.

The effective implementation of the right to choose freely assumes a mature consumer. This is a consumer who through a process of formal or informal education has acquired some degree of understanding of personal and household decision-making and budgeting and who has developed a sense of judgment in making buying decisions. As indicated by cell A1 he will also have some insight into the nature of the market economy and an attendant awareness of consumer rights and responsibilities. His chances of making use of his freedom to choose wisely will be enhanced if he is also aware of how his needs will change over the life cycle and of such basic notions as cost-benefit analysis, discounted cash flow, the economics of information and the value of his time.

To make a choice on any basis other than pure whim or impulse the consumer has to articulate (at least to himself) what his major criteria in buying the product really are. He needs information inputs (B1) to help him define what his requirements in the product should be, given his own set of values and circumstances.¹² Should the consumer be unable or unwilling to digest abstract

information for these purposes special agencies in some countries, such as the Verein für Konsumenteninformation (VKI) in Austria, the Citizens Advice Bureaux in the UK and the Konsumentinstitutet in Sweden may provide him with more concrete—and, by inference, more directive—buying advice. (Should this go as far as quoting brand names we are moving to B2.)

A fairly strong case can be made that consumer sovereignty as an ideal is approachable only in an open market system. Analogously, we may say that consumer policy to be meaningful requires open markets. It seems equally clear by now that the maintenance of open markets requires consumer policy. The open market policy serves the purpose of maintaining viable alternatives in the marketplace. Thus, when it is viewed from the perspective of consumer policy it logically belongs under consumer protection (C1). Here we also find regulation of such practices as high-pressure and door-to-door selling and sundry varieties of deception in trade which obscure or undermine freedom of choice.

Proceeding to the right to be informed we observe the need for generic product and materials information as a logical prelude to the choice among models and brands. This type of information (what will a tape-recorder do that a record-player will not?; what is the difference between nylon and cotton in men's shirts?) in our conceptual scheme falls under the heading of education (A2). The furnishing of data about models and brands is the purpose of information policies implementing the right to be informed, as indicated by cell B2.

There is a broadly felt need in Western countries to safeguard the integrity of product information. In several countries there is mounting pressure --within industry as well as without--towards more informative advertising (C2). Without going into the merits of its policy we may note that the U. S. Federal Trade Commission is a pioneer exponent of the view that advertisers should be prepared to substantiate specific claims made for their products, or to retract false or misleading claims.

Tradeoffs and reenforcement. These are characteristics that apply to both consumer rights and consumer policies. Two things are in a tradeoff relationship when one of them is more or less a substitute for the other (rice and potatoes). They reenforce each other when their joint effect is greater than the total effect obtainable if they are used in isolation (a cake as contrasted to its ingredients). Some examples from consumer rights: if the consumer was informed she can choose wisely (reenforcement). If she has chosen wisely, she will likely not need to complain or to worry about safety (tradeoffs). Similarly, if the pleasure-boat owner about to buy a life vest does not wish to take the time to be informed about the market offerings he may simply get a vest with a safety certificate, but may later find he has to complain about the fabric fading in the sun (tradeoffs). If I have taken time to ferret out my buying criteria I will more likely make an informed choice (reenforcement).

With regard to consumer policies the most dramatic examples of reenforcement occur between consumer education and information. While the reenforcement works both ways, it is far stronger in going from education to information than vice versa. Indeed, one may say that effective consumer information presupposes consumer education. Properly conceived, consumer education provides the citizen with the mental apparatus required to receive and evaluate consumer information. Not only that: we have every reason to expect that well-planned consumer education will provide the stimulus needed to get a ravenous and cumulative information-seeking process going among ever-wider circles of consumers. At present, this type of process is active only among a rather small minority.¹³

Conversely, the most dramatic tradeoff in the consumer policy area occur between consumer information and education on the one hand and protection on the other. This is but natural, in view of the fact that the former policies aim at developing the decision-making capabilities of the consumer, while the emphasis in protection frequently is to substitute the judgment of policy-makers for that of the individual. For instance, a high level of general consumer education would obviate the need for legislation for a period of regret after the signing of door-to-door sales contracts. If there were more consumer information programs, advertising would almost surely be more disciplined--and there would be less concern about the substantiation of product claims.

Policy-makers. This third dimension of Figure 1 was left to the imagination of the reader. In a basic and pervasive sense individual buyers making the myriad day-to-day decisions in the marketplace are the crucial makers of consumer policy. As long as we wish to retain a high degree of consumer sovereignty (and the concomitant open market system) this must be so. Our diagram, however, focuses on organized efforts to educate, inform and protect the consumer. In this view, the policy-makers include consumer organizations, other citizen groups, business, government, educational institutions, and the mass media. In view of the fact that reenforcements and tradeoffs are characteristics of consumer rights and responsibilities as well as of consumer policies, it is hardly surprising to find that the same thing applies to the policy-makers.

To illustrate: from a tradeoff point of view, if business did a better job of informative advertising there would be less need for independent consumer information programs. Similarly, in countries where consumer organizations have not been created (as they have in the US, UK and Benelus countries) to engage in broad-scale consumer information programs, this function has tended to gravitate to government (Scandinavia) or to women's, business and labor groups (Switzerland, France). Reenforcement-wise consumer information can have a multiplier effect if the test reports, labels and quality seals of independent consumer information organizations are used in business promotion programs, reprinted by the press, or discussed on public radio and TV. As regards informative labeling and quality certification programs it would appear that multi-party cooperation between policy-makers is a well-nigh indispensable prerequisite to secure both viability and credibility.

Educational institutions play a role essentially confined to the consumer education area (this is not to say that they are entirely without potential as regards information and protection). Unfortunately, however, one may seriously doubt that schools are doing what they should. Conventional home economics courses in no way meet the agenda implied by the cells in the Education row of Figure 1. For reasons which by now should be abundantly clear, we strongly favor the introduction of obligatory consumer education of the type sketched here at the high school level. We do this even though aware of the fact that there is a tendency afoot to simply push problems to the school system whenever other social institutions have proved unable to solve them. It would seem natural that universities and home economics institutes cooperate with business, consumer groups and governments in the planning of such a curriculum. It is also important that the effort be undertaken from the perspective that while such courses would lay the basis, consumer education, like learning in general, is a life-long process.

Building Trust in the Marketplace

While this is not the place to market a particular brand of consumer policy, a few personal reflections may be permitted in conclusion. Our preference is

on education and information rather than on protection--a preference which may have influenced our choice of illustrations in the preceding discussion. This preference is based on the simple notion that there is no logical end to protective measures, just as there is no logical end to paternalism.¹⁴ Too easily to please our taste, protection invites a kind of censorship under which products will have to pass the test of the bureaucrats instead of the test of the marketplace.

It used to be that the austere rule of caveat emptor defined the place of consumers. Many Western countries are now moving at supersonic speed in the opposite direction, towards caveat venditor (let the seller beware). Open markets with a high degree of consumer sovereignty thrive between the extremes. They are based on trust and on respect for mutual rights and responsibilities. Like all institutions of liberal democracy, trust in the marketplace is a delicate thing that must be fostered with care. This calls for a pluralist approach, for decentralized initiatives in the area of consumer policy. The building of trust also calls for a much greater degree of voluntary cooperation between consumers and producers. On this score there is some reason for optimism: that producer and consumer in the end have more interests in common than in conflict is no more remarkable than the fact that employers and employees do.

Footnotes

1. We are aware of, but not discouraged by the fact that in the end the very notion of "interest" is metaphysical. So is "marginal utility" and "quality"--but they are still indispensable concepts.
2. This is even admitted by as engaged a writer as Jean Meynaud in his Les Consommateurs et le Pouvoir (Etudes de Science Politique 8, Lausanne, 1964), p. 120.
3. Some authorities predict that we may find ourselves saddled with governmentally imposed overall limits on individual consumption before the end of this century for ecological reasons.
4. In the post-industrial society case we would ascribe this explosion of aspirations to a combination of educational and economic affluence paired with a secular radicalization of political life, nowhere more clearly observable than in Sweden. Cf. H. B. Thorelli, "Consumer Information Policy in Sweden--What Can Be Learned?", Journal of Marketing (January 1971), 50-55.
5. Louis L. Stern, "Consumer Protection Via Increased Information," Journal of Marketing (April 1967), pp. 48-52, 49.
6. Why should consumers have to listen to the producer-oriented chatter between taxi drivers and taxi dispatchers? Why should we have to put up with 5 x 12 meter billboards and skyscraper-height gasoline signs along the highways?
7. Whether this "negative responsibility" is better to be looked upon as "right" or "freedom" is a semantic nicety that we fortunately do not have to discuss.
8. In operating an electric saw, for example.

9. Smoking in bed, for example.
10. Using a razorblade for a knife, for example.
11. Some of the negative consumer rights also call for enforcement effort, such as the right not to be subjected to deception.
12. Typical question: what should I look for in a freezer?
13. Hans B. Thorelli, "Concentration of Information Power Among Consumers," Journal of Marketing Research (November 1971), 427-32.
14. We emphasize, however, that the needs of less developed countries and underprivileged minorities elsewhere may call for strong consumer protection measures.

IDENTIFYING DETERMINANTS OF STORE PATRONAGE USING FACTOR ANALYSIS

by
John C. Philpot
Richard C. Reizenstein
and
Daniel J. Sweeney¹

Introduction

Although the selection of a preferred retail store is an important aspect of the consumer purchase decision process, it has received relatively little sophisticated attention in the relevant literature. Conceptually, the consumer's store patronage decision can be viewed as a comparison between certain evaluative criteria and certain perceived retail store characteristics (Engel, Kollat & Blackwell, 1968). The evaluative criteria represent the consumer's desires or expectations regarding various aspects of the retail store. These may be exemplified by store attribute importance variables such as importance of proximity to home, importance of layaway plans, and so forth. The comparison of the perceived characteristics of the retail store image with the consumer's evaluative criteria results in the identification of acceptable and unacceptable stores, indicating those stores which ultimately will and will not be patronized.

From a pragmatic point of view, the retailer needs to identify those evaluative criteria which consumers consider most important and which reflect dimensions of the store's image over which the retailer has some control. The retailer can then alter the characteristics of the store, attempting to make the perceived store image more consistent with the consumers' expectations (evaluative criteria).

Underlying the evaluative criteria are certain customer characteristics including shopping behavior patterns and personal demographic characteristics. These factors may influence the store patronage decision either directly in the form of established shopping habits, or indirectly by influencing the evaluative criteria. Thus, from both a conceptual and a pragmatic point of view, a key element of retail store patronage research is the identification of consistent sets of evaluative criteria used by different groups of customers to judge the acceptability of particular retail stores.

The Current State of Store Patronage Research

Most of the store patronage research currently available can be classified either as store image studies or as retail segmentation studies. The store image studies tend to focus exclusively on the perceived image of specific retail stores and on the differences in perceived store images across consumers (Lazer & Wyckham, 1969; Martineau, 1958; Rich & Portis, 1964; Tillman, 1967). Little attention is given to the relative importance of the various dimensions of perceived store image to the consumer's store

patronage decision. Moreover, there is typically no attempt to associate the store image dimensions with consistent sets of evaluative criteria.

Retail market segmentation studies all too often focus almost exclusively on the demographic and socioeconomic characteristics of retail store customers (Rich & Jain, 1968; Rich & Portis, 1963; Samli, 1970; Thompson, 1967). Rarely is any attempt made to identify retail market segments on the basis of consumers' expectations of retail stores. Exceptions include (Cravens & Cotham, 1970; Hughes, 1966; Sweeney & Reizenstein, 1972).

Some studies have been reported attempting to relate perceived store image dimensions to consumer demographic characteristics (Martineau, 1958; Rich & Jain, 1968; Rich & Portis, 1963 & 1964). Often, however, these analyses are based on cross tabulations of survey data with no indication of the relative strength of the observed relationships.

Furthermore, many patronage studies examine the consumer's choice between alternative stores with such widely differing appeals that the stores often cannot be considered directly competitive (Rich & Jain, 1968; Rich & Portis, 1963). From the retailer's standpoint, the patronage decision in these cases is practically irrelevant.

Finally, retail patronage research has made only sparing use of multivariate analytical tools. Exceptions to this case include (Cravens & Cotham, 1970; Enis & Paul, 1970; Farley, 1968; Hughes, 1966; Sweeney & Reizenstein, 1972). The reliance on analyses that allow only one or two dimensions is a limiting condition on much of the store patronage research noted above.

Focus and Objectives of the Present Study

The present study focused on the store patronage decision of women's apparel purchasers among several directly competing women's specialty and department stores in a medium-sized Southeastern city. The objective of the present research was to identify the major determinants of women's apparel purchasers' store patronage decisions. More specifically, the study was designed to identify systematic relationships among consumers' store attribute expectations (evaluative criteria), their shopping behavior, their demographic characteristics and their preferences for particular women's apparel stores.

Methodology

A total of 500 questionnaires were mailed to a stratified random sample of households in the local retail trading area. One hundred forty-three useable questionnaires were returned, providing data on forty-six individual variables. These variables included:

1. Evaluative Criteria: Relative importance of various store attributes, such as price, personnel, and services, to respondents' store preference and patronage decisions.
2. Customer Characteristics: Respondents' shopping habits and patterns such as frequency of shopping trips and number of stores visited on each trip; and selected demographic characteristics.

3. Respondents' preferences for and allocation of expenditures to specific womens' apparel and department stores.

The store attribute importance variables were rated by respondents using a seven point non-forced choice scale, anchored on a very unimportant to very important basis. All other variables were rated by checking the one of several discrete intervals which most closely reflected that respondent's shopping behavior, demographic characteristics, or store preference. (See further information in the Appendix)

Since the objective of the study was to search for meaningful and stable relationships among the variables, the data were analyzed by means of factor analysis. Using squared multiple correlation coefficients as estimates of communality, principal factors were extracted. Those factors with eigenvalues greater than 1.0 were rotated using the varimax criterion, and a twelve factor model was developed. This model summarized 56.0% of the total variation among the forty-six variables and proved to be highly consistent with a principal-component solution.

Analysis and Interpretation

Factor analysis is a multi-faceted technique. It enables us to express one set of variables in terms of a smaller set of factor variables, and is an approach commonly used for exploratory purposes as well as for tests of hypotheses. Farley, for example, has employed factor analysis in an exploratory manner to identify the determinants of supermarket patronage decision (Farley, 1968). Factor analysis has been employed in this study in a similar context, that is, simply as an exploratory vehicle to determine what types of stable store patronage factors exist.

The dimensions of a given factor can be identified in terms of the variables having heavy loadings on that factor. In addition, the factor axes may, in some cases, represent the clustering of related variables. For example, a set of correlated variables which is not correlated with the remainder of the variables in a study will constitute a separate factor (or subset of factors) in an orthogonal factor analysis of appropriate dimensions. While real data does not meet this condition, a factor may consist of a small set of moderately correlated variables whose members are weakly correlated with most other variables in the study. Such factors have more stability than those consisting of variables that are not directly correlated.

The interpretations that follow make use of these observations. The factor loadings have been examined in conjunction with the original correlation matrix, because correlations between particular pairs of variables loading on a factor are not necessary and cannot be assumed.

Of the twelve factors derived, eight will be examined closely. The remaining four factors, Factors 9, 10, 11, and 12, do not warrant comprehensive treatment as they cannot be clearly interpreted in terms of their component variables.

Factor 1 contains heavy loadings on variables related to age: age of major purchaser of women's apparel, males over 35, females over 35, males

18-35, and females 18-35. Other consumer demographics include family income and length of time as area resident. Shopping behavior variables of distance of store from home and length of time patronizing favorite store are of some importance. Availability of layaway plans is the only store attribute importance variable with a loading greater than .25. Preference for and expenditures in Store A are highly correlated with the dimension of this factor.

Interpretation of this factor suggests that the older purchaser of women's apparel, a loyal patron of her favorite store, a member of an older family, established in the community tends to avoid Store A. This consumer does not find availability of layaway plans an attractive store attribute.

TABLE 1
Factor Loadings for Selected Variables

		Factor 1
	<u>Variables</u>	<u>Factor Loadings</u>
Store Attribute:	Availability of Layaway Plans	-.261
Shopping Behavior:	Distance of Favorite Store from Home	-.266
	Length of Time Patronizing Favorite Store	.388
Consumer Demographics:	Age of Major Purchaser	.690
	Family Income	.341
	Length of Time as Area Resident	.336
	Males, over 35	.768
	Females, over 35	.823
	Males, 18-35	-.663
	Females, 18-35	-.653
Store Preference:	Store A	-.414
Expenditure Allocation:	Store A	-.470

Opposed to this shopper is the younger (18-35 years of age), women's apparel purchaser, newly arrived, less affluent, less store loyal, more oriented toward availability of layaway plans. This latter consumer tends to prefer Store A, a contention supported by the fact that this store does, in fact, deal in sporty clothing with its primary appeal directed at the younger shopper.

Factor 2 comprises heavy loadings primarily on store attribute importance variables identified with a store's proximity to a consumer's home, to other women's apparel stores, and to other retail stores. The consumer demographic variable of female family members under 18 and preference for and expenditure in Store H are both related to this factor.

TABLE 2

Factor Loadings for Selected Variables

Factor 2

	<u>Variables</u>	<u>Factor Loadings</u>
Store Attributes:	Proximity to Home	.653
	Proximity to Other Women's Apparel Stores	.676
	Proximity to Other Retail Stores	.904
Consumer Demographics:	Females, under 18	.296
Store Preference:	Store H	-.251
Expenditure Allocation	Store H	-.353

It is apparent from this factor that shopping convenience, measured in terms of proximity, is an important evaluative criterion for some women's apparel shoppers. One store in particular, Store H, seems to be distinguishable from the others in terms of its failing to satisfy this criterion, as well as in its lack of appeal to younger women. As in the case of Factor 1, this interpretation is generally borne out by reality. Store H is located in a neighborhood shopping center with extremely poor traffic flow and parking. It is relatively isolated in regard to other such centers, and is close to only one other women's apparel store. Its merchandise is mainly oriented toward the more mature woman.

TABLE 3

Factor Loadings for Selected Variables

Factor 3

	<u>Variables</u>	<u>Factor Loadings</u>
Store Preference:	Store E	.963
Expenditure Allocation:	Store E	.742
	Stores Outside of Study	-.327

Factor 3 is composed exclusively of store preference and expenditure variables. Store E has substantial loadings in both areas. Expenditure allocation in stores outside the study is of moderate importance.

The only interpretation which can validly be made here is quite obvious from the above data: Those consumers who prefer Store E purchase a substantial percentage of their women's apparel there. They do not spend a high percentage of their women's apparel dollars in stores outside the study.

Factor 4 loadings are most common on those variables related to income and expenditure. Store attribute importance variables of prices relative to prices of other stores and availability of layaway plans are also important. Moderate loadings are also indicated for expenditure allocation in Stores B, F, and G. Heaviest loadings are on average monthly expenditure on women's apparel and family income. Distance of favorite store from home and length of time as area resident are of some importance.

This factor suggests that consumers with high income who tend to spend a large average amount on women's apparel each month do not consider availability of layaway plans or prices relative to prices of other stores

TABLE 4

Factor Loadings for Selected Variables

Factor 4

	<u>Variables</u>	<u>Factor Loadings</u>
Store Attributes:	Prices Relative to Prices of Other Stores	-.333
	Availability of Layaway Plans	-.353
Shopping Behavior:	Distance of Favorite Store from Home	.346
	Monthly Expenditure on Women's Apparel	.545
Consumer Demographics:	Family Income	.678
	Length of Time as Area Resident	-.254
Expenditure Allocation:	Store B	.253
	Store F	.365
	Store G	.354

to be important evaluative criteria in their store patronage decision. These consumers tend to allocate the greatest percentage of their women's apparel purchases to Stores B, F, and G. Conversely, those consumers who do consider price relative to prices of other stores and availability of

layaway plans as important store attributes would most likely have lower incomes, spend less per month on women's apparel, and would allocate a much lower percentage (if any) of their women's apparel expenditures to Stores B, F, and G. This interpretation is fully supported by the fact that, of the eight specific stores studied, the three noted above are considered the most "exclusive," and carry the highest priced merchandise.

TABLE 5

Factor Loadings for Selected Variables

Factor 5

	<u>Variables</u>	<u>Factor Loadings</u>
Store Attribute:	Availability of Specific Brands	-.290
Shopping Behavior:	Number of Favorite Store Visits per Month	-.699
	Number of Purchases per Month at Favorite Store	-.621
	Monthly Expenditure on Women's Apparel	-.468
Consumer Demographics:	Education of Major Purchaser	.578

Factor 5 seems to focus primarily on variables related to monthly shopping frequency, women's apparel expenditures, and education. Heavy loadings are found on the shopping behavior variables related to the number of times the favorite store is visited monthly, the number of purchases made monthly at the favorite store, and the average monthly expenditure on women's apparel. The consumer demographic variable of educational level of the major purchaser of women's apparel is also heavily loaded. Finally, the store attribute importance variable of availability of specific brands is of some relevance, though its loading is much lower than those previously discussed.

This factor seems to indicate that those who shop and make purchases frequently in their favorite store, consequently spending a high average amount of money per month on women's apparel, tend to have a relatively low educational level compared to the remainder of the sample. These individuals seem (though the relationship is weak) to find availability of specific brands important, providing a possible insight into a reason for such frequent patronage of the favorite store. It also appears that a subset of those who do not shop or purchase frequently at their favorite store, and who do not spend a high average monthly sum on women's apparel, are relatively well educated. This segment does not seem to be as brand conscious as their lesser educated counterparts.

TABLE 6
Factor Loadings for Selected Variables

		Factor 6
	<u>Variables</u>	<u>Factor Loadings</u>
Store Attribute:	Availability of Specific Brands	.260
Shopping Behavior:	Distance of Favorite Store from Home	-.473
Store Preference	Store B	.843
Expenditure Allocation:	Store B	.803

Factor 6 seems to be a measure of the appeal of Store B contrasted with distance of the favorite store to home. The heaviest loadings are those on preference for and expenditure allocation to Store B. The shopping behavior variable of distance of favorite store from home is also quite noteworthy, with the store attribute importance variable of availability of specific brands being of some, though weaker importance.

This factor appears to indicate that part of the unique appeal of Store B is that distance from the consumer's home to the store is short. This is supported by the fact that Store B is one of the most exclusive specialty shops, and is located within five minutes of one of the most exclusive residential areas of the city. It can also be noted that it appears (though the relationship is weaker) that patrons of Store B tend to prefer specific brands.

TABLE 7
Factor Loadings for Selected Variables

		Factor 7
	<u>Variables</u>	<u>Factor Loadings</u>
Store Attributes:	Range of Sizes Available	-.427
	Availability of Charge Accounts	.299
Consumer Demographics:	Age of Major Purchaser	-.383
	Males, under 18	.519
	Females, under 18	.651
	Children at Home	.888

Factor 7 is characterized by variables related to young families with children. Strong factor loadings may be observed for the following con-

sumer demographic variables: (1) Children at home; (2) Males, under 18; (3) Females, under 18; and (4) Age of major purchaser of women's apparel (though the loading here is weaker than in the previous three variables). The store attribute importance variable of range of sizes available was also considered quite important, as was availability of charge accounts.

This factor certainly seems to revolve around younger families, with children under 18 at home. These consumers do not tend to place much emphasis on the range of sizes available as an evaluative criterion for women's apparel. They do, however, seem to indicate the importance of the availability of charge accounts.

TABLE 8

Factor Loadings for Selected Variables

Factor 8

	<u>Variables</u>	<u>Factor Loadings</u>
Shopping Behavior:	Number of Purchases per Month at Favorite Store	-.296
	Distance of Favorite Store from Home	-.267
Store Preference:	Store D	-.825
Expenditure Allocation:	Store D	-.877

Factor 8 contains extremely heavy loadings on Store D in regard to both store preference and expenditure allocation variables. Shopping behavior variables of number of purchases per month at favorite store and distance of favorite store from home are of some importance, though they are considerably less relevant than the variables directly relating to Store D.

The loadings for this factor indicate the overwhelming predominance of Store D in both store preference and expenditure allocation to favorite store. In addition, Store D customers tend to make more purchases per month and to travel a greater distance from home than those who do not patronize Store D. This interpretation is strongly supported by the fact that Store D is, by far, the dominant department store in the city, thus indicating that more people would travel further to patronize it. Moreover, the much wider variety of general merchandise carried in the department store would suggest a higher purchase frequency in Store D than in a store specializing exclusively in women's apparel.

Conclusions

Several general conclusions can be drawn from the study regarding the determinants of retail store patronage decisions, and the structure and methodology of future store patronage research.

1. The study identifies several sets of variables which appear to be related to the store patronage decision. In some cases the important variables relate to store attribute importance measures, in other cases to shopping behavior characteristics or to consumer demographic characteristics. The store attribute importance measures are particularly vital since these form the basis for identifying the evaluative criteria used by the consumer in selecting a preferred retail store. Having isolated the relevant evaluative criteria, the retailer should be able to redesign certain aspects of his store image to become more consistent with the consumer's store patronage expectations.
2. It appears that significant progress can be made in defining the determinants of store patronage if the analyst focuses on the interrelationships among a variety of different types of consumer characteristic and store attribute variables, using analytical tools capable of reflecting the multiple dimensions of the store patronage decision process. The present study indicates the value of one such multivariate technique, factor analysis, in the study of retail store patronage decisions.
3. The present study suggests that the determinants of the store patronage decision are somewhat unique to stable groups of consumers as well as to specific stores, even though the stores are intensely competitive. This suggests that a highly differentiated market segmentation strategy might be very effective in the retail women's apparel market.
4. Additional research should be directed at identifying a more complete and sensitive set of store attribute and consumer characteristic variables for store patronage research. This need is underscored by the fact that five of the twelve factors identified in this study had heavy loadings on unique stores but only scattered loadings of moderate value on store attribute importance, shopping behavior, and consumer demographic variables.

Footnotes

1. John C. Philpot is Assistant Professor of Statistics. Richard C. Reizenstein and Daniel J. Sweeney are Assistant Professors of Marketing. All are of the University of Tennessee.

References

- Cravens, David W., & Cotham, James C. Identifying Market Segments Using Canonical Correlation Analysis. A paper presented at the American Marketing Association 1970 Fall Educators Conference, Boston, Massachusetts, September, 1970.
- Engel, James F., Kollat, David T., & Blackwell, Roger D. Consumer Behavior. New York: Holt, Rinehart, and Winston, Inc., 1968.
- Enis, Ben M., & Paul, Gordon W. 'Store Loyalty' as a Basis for Market Segmentation. Journal of Retailing, 1970, 46, 42-56.
- Farley, John U. Dimensions of Supermarket Choice Patterns. Journal of Marketing Research, 1968, 5, 206-208.
- Harman, Harry H. Modern Factor Analysis. Chicago: The University of Chicago Press, 1967.
- Hughes, David G. Developing Marketing Strategy through Multiple Regression. Journal of Marketing Research, 1966, 3, 412-415.
- Lazer, William & Wyckham, Robert G. Perceptual Segmentation of Department Store Markets. Journal of Retailing, 1969, 45, 3-14.
- Martineau, Pierre. The Personality of the Retail Store. Harvard Business Review, 1958, 36, 47-55.
- Rich, Stuart U. & Jain, Subhash C. Social Class and Life Cycle as Predictors of Shopping Behavior. Journal of Marketing Research, 1968, 5, 41-49.
- Rich, Stuart U., & Portis, Bernard. Clues for Action from Shopper Preferences. Harvard Business Review, 1963, 41, 132-149.
- Rich, Stuart U. & Portis, Bernard D. The Imageries of Department Stores. Journal of Marketing, 1964, 28, 10-15.
- Samli, A. Coskun. Interrelationship between the Market Segments and the Buyer Behavior. Talk delivered before the X ESOMAR Seminar held at Lucerne, November 2-5, 1969, reprinted in David J. Rachman, Retail Management Strategy: Selected Readings. Englewood Cliffs: Prentice-Hall Inc., 1970, 97-112.
- Sweeney, Daniel J. & Reizenstein, Richard C. Developing Retail Market Segmentation Strategy for a Women's Speciality Store Using Multiple Discriminant Analysis. A paper presented at the Fall Educator's Conference, The American Marketing Association, August, 1972.
- Thompson, B. An Analysis of Supermarket Shopping Habits. Journal of Retailing, 1967, 43, pp. 17-29.
- Tillman, Rollie. Semantic Differential Measurements of Consumer Images of Retail Stores. The Southern Journal of Business, 1967, 2, 67-73.

Appendix

Scales of Measurement for VariablesStore Attributes:

All variables (prices, brands, range of sizes, etc.) were on a 1-7 rank-order scale, where

1 = attribute very unimportant to the consumer

7 = attribute very important to the consumer

Shopping Behavior

All variables were on the positive integer scale, with integers corresponding to the following bases:

<u>Variable Name</u>	<u>Basis for Scale *</u>
No. of Favorite Store Visits per Month	The Variable Itself
No. of Purchases/Mo. at Favorite Store	The Variable Itself
Distance of Favorite Store from Home	2 Mile Intervals
Length of Time Patronizing Favorite Store	3 Month Intervals
Monthly Expenditure of Women's Apparel	\$20 Intervals

*Unless otherwise specified, there was some compression of variable on the scale for the upper and/or lower values of the variable.

Consumer Demographics:

The following variables were on the positive integer scale, corresponding to the following bases:

<u>Variable Name</u>	<u>Basis for Scale</u>
Age of Major Purchaser	10 Year Intervals
Education of Major Purchaser	The following educational divisions: 1 = grade school; 2 = junior high; 3 = high school; 4 = college (2 yrs); 5 = college (4 yrs); 6 = grad. school
Family Income	\$5000 Intervals
Length of Time as Area Resident	3 Month Intervals

The remaining variables (Male over 35, Female over 35, Male 18-35, etc.) took on either 0 or 1 values, where:

- 0 = there was no member of the family corresponding to the age group
- 1 = there was at least one member of the family corresponding to the age groups

Store Preference:

All variables (Store A, Store B, etc.) indicated consumer rankings of the respective stores on a -20 to -1 scale, where the largest value (-1) indicated the consumer's favorite store; the second largest value (-2) indicated the next most favored store, etc. This scaling allowed the store preference variables to have, in general, the same signature for a given store as the corresponding expenditure allocation variables.

Expenditure Allocation:

All variable (Store A, Store B, etc.) consisted of the consumers' annual expenditure for women's apparel on a percentage scale, from 0 to 100.

RESOLVING FACTOR STRUCTURE DISTORTIONS¹ IN CONSUMER RESEARCH

Robert H. Taylor, E. Laird Landon, Jr.,
and Jerome E. Scott²
University of Colorado²

Factor analysis has long been one of the principle tools of behavioral research. Its uses range from the reduction of large amounts of data to manageable proportions, to lending insights and understanding to complex behavioral phenomena. Because of this long history of use and frequent appearance in both psychological and marketing literature, there is a feeling of confidence in the use of the technique, when in fact there are many hidden and unresolved problems associated with the use of factor analysis.

The purpose of this paper is to investigate one potential source of trouble relevant to consumer research and to suggest some possible solution procedures. Consideration of this problem should lead to more valid inferences being drawn from behavioral research.

One of the major uses of factor analysis is to derive the underlying structure, or factors, of attitudes or behaviors by analyzing the inter-relationships between variables (test scores, overt behavior, etc.) measured on a sample of individuals. An implicit assumption in this approach is that all individuals are guided by the same latent relationships: that is, there is an assumption of a 'universal' factor structure where all individuals perceive the same underlying domain.

Unfortunately, there may be no a priori reason for making this assumption. In fact, the opposite may be more tenable and may be exactly what the researcher is trying to find out. Therefore, to use pooled data (a factor analysis conducted across the total sample) when there are fundamental differences between subsets of subjects can lead to misinterpretations concerning the underlying factors. A factor analysis performed on the total group may differ radically from the analysis performed on the sub-groups of which it is comprised.

The distortions which can result are illustrated in Figure I. The sample consists of two well defined sub-groups, but factoring the total data set would produce one set of axes, while factoring of the two sub-groups separately would produce two entirely different sets of axes (the primed factors).

If the purpose of the analysis is data reduction, then the factor scores are of interest. In this case the analysis would be performed on the total group in order to maintain a common reference point for the factor scores. However, if one is interested in the underlying structure of behavior, the groups should be analyzed separately. To pool the data might lead to a non-meaningful factor matrix.

The technical problem is one of defining modes in the data set, and the subsequent examination of the sub-structures derived from factoring. This problem has been recognized in the clustering literature (Tryon and Bailey, 1970) and in scaling (Tucker and Messick, 1963) as well as the literature dealing with factoring (Harman, 1967), but has not generally been applied to empirical studies dealing with consumer behavior.

The first step in the analysis is to separate the data base into sub-groups which are of interest to the researcher. Such groupings are made from hypotheses concerning differential response patterns. Typical examples would be examination of males versus females, blacks versus whites, city versus country residents, etc. Each sub-group would then be separately factor analyzed, in addition to a factor analysis on the total group (to be used for com-

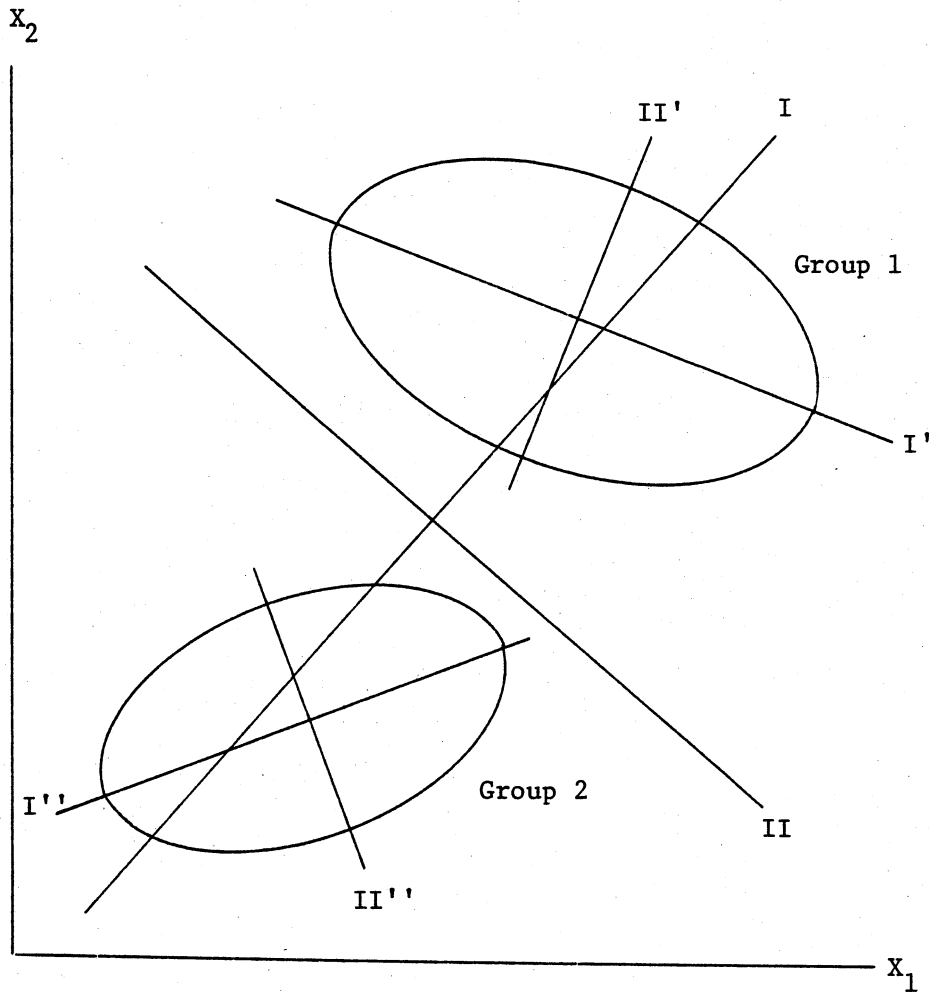


Figure I

Pooled Versus Separate Group Factoring

parative purposes). The second step required in the analysis is to compare and analyze the resulting factor solutions to determine the degree of similarity between them.

Methods for Comparing Factor Solutions

The problem of factor matching involves two aspects of factor structure: the analysis of patterns of loadings, and the analysis of the level of loadings. Three approaches have been proposed for comparing factor structures: visual inspection, vector by vector methods, and total matrix comparisons.

Visual Approaches. For the researcher with a good understanding of the problem being investigated, a simple visual inspection of the factor structure might be sufficient to determine the degree of similarity or difference. In addition to level and pattern, the complexity of variables and the communalities can also be examined. The personal impressions formed by such inspection can often lead to subtle insights which raise

questions about the comparability of solutions. A visual inspection is generally recommended as a preliminary step in the analysis and may be followed by one or more of the mathematical techniques available.

Vector Comparisons. The vector by vector approaches focus on how congruent any two factors, considered separately, seem to be. Harman (1967) proposed a root mean square approach for determining the extent of agreement. Using his notation:

$$(p - q)_{\text{rms}} = \sqrt{\sum_{j=1}^n (1^{a_{jp}} - 2^{a_{jq}})^2 / n}$$

gives the agreement between Factor p of the first study and Factor q of the second study across the n common variables. The a 's represent the factor loadings for the j 'th variable on the p 'th factor, etc. This measure is proportional to the distance between factors (positioned in the common n -dimensional space) and imposes both pattern and level restrictions on the definition of similarity. In addition, the index may be difficult to interpret since all one knows is that a perfect fit would yield an index of zero.

A second possible index is the coefficient of congruence, defined as:

$$\phi = \frac{\sum_{j=1}^n 1^{a_{jp}} \cdot 2^{a_{jq}}}{\sqrt{\sum_{j=1}^n 1^{a_{jp}^2} \cdot \sum_{j=1}^n 2^{a_{jq}^2}}}$$

with the same meaning and variables as the above formula. This coefficient is the cosine of the angle between factors, and varies from zero (non-correspondence) to \pm one. It is less restrictive since the lengths of the vectors are normalized prior to comparison. Thus, level differences are partially removed.

Additionally, the simple correlation coefficient between factor loadings could be used to measure similarity. This is least restrictive in its requirements since it is a measure of pattern without regard for the magnitude of the loadings.

Several cautions are important to observe when employing these coefficients. First, the researcher may be comparing factor structures from two different studies where only a sub-set of the variables are common to both. In this case, the indexes are applicable only to the common variables. Cattell (1952) points out that this can lead to mistaken judgements about the similarity of the factors if the elimination of dissimilar variables also removes the variables which load heavily on the factors and are therefore the defining variables. The comparisons may then be using only the variables with modest loadings which can lead to assertions that the factors are similar when they are defining different constructs (or that they are different when in fact they are defining the same construct). To reduce the likelihood of such spurious judgements, Cattell (1952) suggests that only variables which uniquely define a factor be included in the comparison, and that communalities provide the basis for selecting these variables. If one restricts comparisons to only those variables with loadings exceeding .8; then (for orthogonal factors) its loadings must be low on the remaining factors.

A further problem with the coefficient of congruence is that the relative size of the index tends to be quite large, even for factors which are

quite different. Penneau and Newhouse (1964) point out that the lowest value the index can take for comparisons involving factors with only positive loadings is .67, and that factors with similar patterns but quite different magnitudes will result in indices of above .9.

Matrix Comparisons. The disadvantage of the vector approaches is that the factor loadings are taken as given by the rotation method. However, exogenous influences may affect independent rotations and confound comparisons. Since factor rotations are arbitrary (an infinite number of solutions exist), a better approach might be to rotate one of the factor solutions to a least squares fit with the other. Ahmavaara (1954) developed such a target matrix approach.

A difficulty with Ahmavaara's approach is that the solution depends upon which matrix is the target and which is the rotated factor loading matrix. In general, there may be no a priori reason for assigning a particular loading matrix as the target. In this situation it is better to rotate both factor solutions simultaneously to a common point of maximum congruence. A procedure was developed by Horst (1965) as a generalization of the two group canonical correlation problem where factor loadings are orthogonally rotated so that the correlation between loadings is maximized. Since this procedure is based on the correlation between factor loadings, it is a measure of the pattern similarity of loadings and does not measure differences in level.

Results of Factor Structure Comparisons

In a recent study, several hypotheses were generated concerning differential behavior of high and low need for achievement subjects in Bayesian probability revision experiments. However, a key to this relationship is that risk preference is considered a major component of need for achievement (Feather, 1959), and this relationship must be demonstrated.

Ninety-seven males were measured on a twenty-six item measure of need for achievement (Mehrabian, 1968) and eight risk preference items. The pooled group was factor analyzed to determine the dimensionality and to find the key definers of the dimensions. The results are shown in Table I. The most surprising outcome was the lack of a relationship between the risk items (which emerged as an independent factor) and need for achievement.

The data were separated into two groups on the basis of need for achievement and each group was separately factored. Two problems arise with this procedure. First, the variables used to define the groups were used again in the factor analysis. If the intent were to predict membership, this procedure would surely bias the results. However, since relationships between variables were the focus, the procedure was felt to be valid. Second, the splitting procedure reduced the sample size for each group to forty-eight. Since thirty-four variables were being used, the stability of the correlation coefficients are highly suspect (Cattell (1952) suggests four times as many cases as variables).

The results of the separate factor analyses are shown in Tables II and III. In both solutions, risk preference items loaded on factors containing need for achievement items, indicating a relationship between these two sets of variables within the groups. A visual inspection of which variables define the factors for each of the sub-groups strongly suggests that the factors appear to be different constructs between groups. Further, both sub-group solutions are different from the total group solution. It would appear that risk preference is related to and interacts with need for achievement, and that this relationship is hidden by the total group factoring.

Table I

Total Group Factor Structure

Factor					
<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>
12	27	17	15	11	8
13	29	18	23	20	9
	30	25		24	26
	31				
	32				

Note: Entries correspond to variable numbers which are key definers of the dimensions. Variable numbers 1 to 26 are need for achievement items, and 27 to 34 are risk items.

Table II

Low Need for Achievement Group Factor Structure

Factor					
<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>
5	3	12	28	14	7
11	15	13	31	17	24
16	18		34	25	26
18	23			26	
21				27	
32					

Note: Entries correspond to variable numbers which are key definers of the dimensions. Variable numbers 1 to 26 are need for achievement items, and 27 to 34 are risk items.

The sub-group factor solutions were then compared factor by factor using the coefficient of congruence (Table IV). An inspection of the coefficients indicates very low similarity for any pair of factors between sub-groups. These results are consistent with the visual examination.

The sub-group factor solutions were then analyzed using Horst's matrix comparison approach to obtain a simultaneous comparison of the factor structure (Table V). This procedure attempts to maximize the correlation between loadings. The largest correlation is .77 which indicates a relatively low degree of correspondence between the most similar factors of the two structures. The remaining correlations fall quite rapidly with virtually no relationship

indicated for the least similar pair. This procedure does indicate somewhat higher correspondence for the pairs of factors than the coefficient of congruence, but this is a result of the rotation of the structure to maximum congruence prior to the calculation of correlations. This has the effect of removing part of the angular separation of the factors, while the use of correlations between factors removes level differences.

Table III

High Need for Achievement Groups Factor Structure

		Factor					
		<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>
	18		3	9	5	6	2
	27		4	34	7	14	11
	29		8		12	28	19
	30		24				20
	31						24
	33						

Note: Entries correspond to variable numbers which are key definers of the dimensions. Variable numbers 1 to 26 are need for achievement items, and 27 to 34 are risk items.

Table IV

Factor Comparisons Coefficient of Congruence

		High Groups Factor					
		1	2	3	4	5	6
	1	-.05879	.11184	-.11306	-.13673	.05765	.40694
	2	.20226	-.13695	-.46337	-.07688	-.12246	-.09346
Low	3	.15137	.10034	-.01790	-.28245	.38077	.12897
Group	4	-.44045	-.05740	-.14169	-.17014	-.26469	.15715
Factor	5	-.01391	-.00995	-.04374	.16792	.10793	.04769
	6	.23527	.14070	.04773	.08536	.14638	.24375

The guidance given to the Bayesian research is to verify that there is a relationship between risk preference and need for achievement, but that the direction and extent of that relationship is a function of the degree of need for achievement. These results also indicate that a simple pooled group factoring may hide the true structure of the system because of an invalid assumption of a universal factor structure.

Table V
Factor Comparisons
Horst's Maximal Congruence

.77				
	.66			
		.52	(0)	
			.41	
(0)				.20
				.03

Footnotes

1. The authors would like to thank the University of Colorado Computing Center for providing computing time.
2. The authors are all Assistant Professors of Marketing at the University of Colorado.
3. All factor analyses were conducted using the squared multiple correlation coefficient as the communality estimate. Varimax rotations were performed. Since the total group structure included six factors (eigen-values greater than one), all subsequent analyses were carried to six factors.

References

- Ahmavaara, Y. The Mathematical Theory of Factorial Invariance Under Selection. Psychometrika, 1954, 19.
- Cattell, R. B. Factor Analysis. New York: Harper and Row, 1952.
- Harman, H. H. Modern Factor Analysis, 2nd ed., rev. Chicago: The University of Chicago Press, 1967.
- Horst, P. Factor Analysis of Data Matrices. New York: Holt, Rinehart, and Winston, Inc., 1965.
- Mehrabian, A. Male and Female Scales of the Tendency to Achieve. Educational and Psychological Measurement, 1968, 28, 493-502.
- Pinneau, S. R. & Newhouse, A. Measures of Invariance and Comparability in Factor Analysis for Fixed Variables. Psychometrika, 1964, 29, 271-281.
- Rummel, R. J. Applied Factor Analysis. Evanston, Illinois: Northwestern University Press, 1970.
- Tryon, R. C. & Bailey, D. E. Cluster Analysis. New York: McGraw-Hill, 1970.
- Tucker, L. R. & Messick, S. An Individual Difference Model for Multi-dimensional Scaling. Psychometrika, 1963, 28, 333-367.

BACKWARDS SEGMENTATION USING HIERARCHICAL CLUSTERING
AND Q FACTOR ANALYSIS

Flemming Hansen
Marketing Director
T. Bak-Jensen A/S/A.I.M., Copenhagen¹

The Problem

The idea behind market segmentation is simple. If you can divide a larger market into smaller segments with different preferences and subsequently adjust your product to the preferences in the different segments, then you reduce the overall distance between what you are offering to the market and what the market requires. By doing so the marketer improves his competitive position.

For market segmentation to be successful, however, two requirements must be met:

1. It must be possible to identify market segments with different behavioral and/or preferential patterns.
2. It must be possible to reach the different segments with differential media, distribution, product, communication message or other strategies.

Although the extent to which the second condition is met may vary from situation to situation, some possibilities will almost always exist when first the segments are identified. Consequently in the following the concern is primarily with the question which has become crucial to most marketers, namely how can different segments be identified?

The question as to the extent to which useful segments can be constructed can be divided into two. First, it must be possible to identify the different segments, secondly--for a differentiated marketing effort to be worthwhile--the segments must respond differently, that is the segments must be characterized by different demand elasticities. The latter, however, can be expected to be the case when segments can be identified which behave differently or which differ in the consumer values governing the behavior.

Therefore it is a major problem what kind of variables to use in an attempt to segment consumers. It is extremely important that the variables chosen really do relate to the behavior of the consumers. In the following pages some of the most commonly used variables are reviewed and following this two studies using different types of variables are reported.

Segmentation Variables

It may be possible to explain consumer behavior by means of personality variables. In the following, the application of personality tests in studies of consumer behavior will be reviewed. The discussion covers studies where larger tests have been used as well as studies where single personality dimensions are explored.

Systematic differences in consumer behavior may be studied based upon socioeconomic and demographic variables also. This will be explored, and in this connection consumer types such as innovators and opinion leaders are discussed.

It is possible that systematic differences in the consumer's self-image may explain differences in consumer behavior. Therefore, studies dealing with the relationship between self-image and product images (image congruence) will be reviewed also.

It may be possible to infer consumer values from observation of behavioral differences. That is, based upon systematic differences in consumer behavior, it may be possible to segment consumers meaningfully. Reports from studies which have applied this kind of "backward segmentation" are reviewed also.

Finally, it may be possible to learn about consumer values simply by asking about the interests, likes and values of the consumer. Also such studies are discussed.

1. The Use of Personality Inventories

In studies of consumer behavior, the most commonly applied personality inventory is "Edwards' (1954) Personal Preference Schedule" (EPPS). This inventory, which builds upon Murray's (1938) study of personality and motivation, is a forced choice, paper and pencil test classifying individuals along 15 personality dimensions. The classical study is reported by Evans (1959) and a replication is reported in Evans (1968). Evans administered selected items from the EPPS to Ford and Chevrolet owners and attempted to predict brand choices based on a discriminant analysis where the personality scores were used as independent variables. With this it was possible to classify 62.9 percent of the owners correctly; a prediction which was only slightly improved when 8 demographic variables were added to the personality variables.

Other authors (Kuehn 1963 & Marcus 1965) have re-analyzed the data and improved the predictions only slightly and other applications of EPPS are reported by Koponen (1960), and by "The Advertising Research Foundation (1964)". These studies deal with grocery products and attempt to predict brand choice, brand loyalty, quantity purchased, and other aspects of consumer choices. Neither of the studies have been very successful. In line with this are the findings reported by Massy, Frank and Lodahl (1968), who tried to explain differences in a number of consumer purchase variables.

Among all the attempts with the EPPS only Claycamp (1965) reports moderately successful results from a study of thrift deposit owners in commercial banks and in saving and loan associations.

Many other personality inventories have been applied. Westfall (1962) was moderately successful with "Thurstone's Temperament Schedule" (Thurstone, 1953) in an automobile study. Similarly, Kernan's (1968) application of the "Gordon Personal Profile" has not produced convincing results, and with the same test Tucker and Painter (1961) found only small correlations between personality types use of studied products (such as headache remedies, vitamins, mouthwash, cigarettes, chewing gum, alcoholic drinks, fashions, deodorants and automobiles). Other studies with only modest success are repeated by Wicks and Nelson (1967), who used the "Guilford-Zimmerman Temperament Survey"; and by Ruch (1966), who analyzed heavy and light users, and loyal and non-loyal users of grocery products, with "McCloskey Personality Inventory". Similarly Myers (1967) found that consumers' attitudes towards private brands are not significantly related to their score on "Cattell's 16 Personality Factor Inventory", and Brim, et al. (1962) were unable to explain differences in choice process behavior with personality scores obtained with "Taylor's Manifest Anxiety Scale". In line with the results are also those reported by Robertson and

Myers (1969) who found no correlations between, on the one hand measures of opinion leadership and innovative buying behavior and on the other hand personality scores obtained with the "California Psychological Inventory". Similarly, Rizzo and Naylor (1964) who applied Allport et al.'s (1960) "Study of Values" report negative results, and so does Scott (1957) who applied the "Minnesota Multiphasic Personality Inventory" in an attempt to predict motion picture preferences.

Altogether, personality inventories have not proven very useful in studies of consumer choice behavior: Wells (1966) summarizes the results nicely: "The findings of these studies have been very consistent. Almost always they have resulted in statistically significant correlations that have been too small to be of much practical value." (p. 187). It would be tempting to conclude that systematic differences in the values which consumers hold do not relate to the choices they make. However, before this view can be adapted, two alternative explanations must be rejected. First, personality tests are highly complex and difficult tests to construct, and those presently available may not be ideal. Secondly, the tests which have been applied have been developed for clinical and other uses, and it is questionable whether they can be expected to explain differences in a completely unrelated area such as consumer choice behavior. To the present author, the last explanation seems the most plausible. As Steiner (1966) suggests "I do not blame psychologists for the failure reported here in attempts to explain behavior by using certain tests of personality--you cannot take just any tool off the shelf simply because it happens to be there and expect that it will be the best tool for your job." (p. 208).

2. Individual Personality Traits

Some authors, instead of using general personality inventories, have tried to relate consumer behavior to individual personality traits. Some of the more promising attempts have dealt with inner-other directedness, self-esteem and self-confidence, propensity to take risk, achievement motivation, and measures of cognitive style.

Inner-other directedness reflect the individual's tendency to rely upon others (other directedness) in decision making and evaluation of information, as opposed to a tendency to rely upon own judgements and values (inner directedness). The dimension relates to Riesman's (1961) distinction among traditional-, inner- and other-directedness, and it has normally been measured by means of a 36 item social preference scale constructed by W. Kassarian (1962). With this H. Kassarian (1965) found inner- and other-directedness to be related to consumers' preferences for persuasive communication and similarly, Arndt (1968) and Donally (1970) found it to relate to innovativeness.

"Self-esteem" or "generalized self-confidence" (Cox and Bauer, 1964) is a personality dimension which reflects the individual's feelings of social adequacy and of confidence in their own ability to cope with problems and aggressions.² Generalized self-confidence has frequently been measured with some or all of 23 items constructed by Janis and Field in Hovland and Janis (1959). Generally it has been related to individuals' susceptibility to persuasive communication.

However, Schuchman and Perry (1969) question the validity of several of the findings and only slight support is presented in the findings reported

by Arndt (1967 and 1968). And, also, Ostlund (1969) was unable to establish any significant relation between generalized self-confidence and innovativeness.

Individual differences in willingness to accept risk are reported by Brim and Hoff (1957) and by Cunningham (1967b). The latter found that subjects who tend to perceive high risk in one product area also tend to perceive high risk in other product areas, and reverse, subjects who perceive little risk in one product area also are more likely to perceive little risk in other product areas.

Several studies confirm the importance of perceived risk as a personality variable. Arndt (1967 and 1968) reports that low risk perceivers are more likely to be innovative, and Cunningham (1964) reports that the extent to which consumers engage in personal communication in connection with grocery products is related to their tendency to perceive risk. Furthermore a number of studies have indicated relationships between loyalty and perceived risk (Cunningham, 1967a; Arndt, 1967; and Brody & Cunningham, 1968), and between innovativeness and a propensity to perceive risk (Arndt, 1968 and Ostlund, 1969).

Achievement motivation has been measured in several different ways. Normally, however, it is inferred from projective measures, most commonly the Thematic Apperception Test (TAT).

It is a basic proposition in the theory of the achievement motive that economic activity is related to achievement. In line with this, Morgan (1966) reports significant differences in income and in spending behavior depending upon achievement motivation. Similarly Boulding (1960) relates differences in consumer behavior to two personality types characterized as "integrated achievers" and "satisfied securers". Moreover, a number of studies have found need achievement to be related to risk taking propensities. For example, Scodel, et al., (1959) conducted a study where they applied a large number of different personality variables, but found only achievement motivation to be significantly related to risk-taking.

A major problem with this research has been that the different need achievement measures do not correlate well, and that no attempts have been made to relate achievement motivation to consumer choices. However, if the measurements problems can be overcome this variable may prove to be of some use in studies of consumer behavior.

Cognitive style reflects the way in which the individual approaches problems: whether they apply more or less wide categories, whether they place major emphasis on problem solving goals or upon social goals, and whether they strive for cognitive clarity or simplicity.

A couple of studies have dealt with these variables. Popielarz (1967) reports that consumers who use wide conceptual categories are more willing to accept new brands, and Phares and Davis (1966) found them to make larger adjustments in expectations following disconfirming experiences. Similarly, Cox (1967b) found that subjects with a high need for cognitive clarity are more susceptible to persuasive influence, and in the same study it also appeared that subjects who could be characterized as clarifiers (those who tend to clarify an issue) as opposed to simplifiers (those who tend to simplify an issue) respond differently to persuasive communication. Finally, Wilding and Bauer (1968) found subjects with predominantly social goals to react significantly different to communication as compared with subjects with predominantly problem solving goals.

Taken together the findings suggest that cognitive style may be an important variable, but so far the interrelations among the different measures is completely unexplored; and not until we have a better understanding of the nature of cognitive styles can more general hypotheses be formulated.

In contrast with the traits measured in standardized personality inventories, most of the variables discussed here have emerged in studies of consumer and similar behavior. As shown, there is some evidence in favor of such variables. However, many questions remain to be answered, and the findings which have been reported do not suggest simple relationships between personality variables of this kind and the consumers' choices and values.

3. Socioeconomic and Demographic Variables

The previous discussion suggests that so far personality variables have not proven to be highly useful for the purpose of market segmentation. This would be less important if segmentation could be accomplished with socioeconomic and demographic variables; and of course, to some extent, these variables are useful. For example, it is obvious that homeowners are more likely to purchase outdoors paint and that households with babies are more likely to purchase baby food, etc. That is, socioeconomic and demographic variables can be used to define that segment of the total population which can possibly demand a certain product. However, finer discriminations can rarely be made with these variables. Even though some marked differences may exist among the users of different brands, and retail stores may attract different consumers, it has often been found that the ability of socioeconomic and demographic criteria to discriminate among consumers is relatively limited. As mentioned earlier, Evans (1959) only slightly improved the discriminative power of his equation which he used to predict ownership of Chevrolet and Ford when he introduced demographic variables, and Frank et al. (1967) found only slight socioeconomic and demographic differences among consumers who purchased more or less expensive grocery items. Similarly, Frank and Boyd (1965) and Myers (1967) found no differences among consumers who prefer private brands as compared with those who prefer manufacturers' brands, and Kuehn (1966), Frank (1967a), and Massy, et al. (1968) report that loyal consumers cannot be identified by means of demographic and socioeconomic characteristics. Finally, Frank (1967b) reports that heavy versus light buyers of a grocery product do not have different socioeconomic characteristics. Altogether, the available evidence warrants the conclusion by Frank (1968): "For the most part, socioeconomic characteristics are not particularly effective bases for segmentation" (p. 53).

4. Social Class

An individual's social class reflects the way in which he is perceived by others in the society. Warner and Lunt (1941) suggest that a social class consist "of people who are believed to be, and are accordingly ranked by the members of the community, in socially superior and inferior positions" (p. 82).

Several studies have attempted to relate social class to consumer behavior. Graham (1956) found different adoption patterns in different social classes for products such as television, canasta, super-markets, and medical services. Similarly, Martineau (1958) reports that many aspects of spending behavior and of store choices are related to social class. However, Brim, et al. (1962) found social class to have only little influence upon decision process variables and with regard to consumer behavior Rotzoll (1967) suggests that finer distinctions among social classes are of doubtful value. It is rarely possible to identify more than two separate classes. Also in the early sixties

Martineau (1963) found many social class differences in shopping behavior to be disappearing and Rich and Jain (1968) after reviewing the literature, suggest that with the rapid changes which occur in income, leisure time, education and the movement to suburbia, social class differences which may have existed earlier, are likely to disappear in the future.

5. Family Life Cycle

Consumers in different stages of the family life cycle are expected to behave differently, and possibly they have differing value structures. Often significant differences in spending behavior, savings, and possession of different durable products have been reported. However, no studies have found important variations in brand choices or in connection with purchases of non-durable products.

As with social class, Rich and Jain (1968) suggest that many differences traditionally associated with the family life-cycle tend to be ruled out by other changing factors in contemporary societies.

6. Innovators and Early Adopters

Much effort has been expended in attempts to identify consumers who are likely to accept new products early. To the extent that such consumers can be defined as a special market segment, they represent a group of consumers of particular importance.

Two questions must be raised in connection with innovators as a special market segment. First, do those who adopt a particular product early differ from those who adopt it later? Secondly, are those consumers who adopt early in one product area also likely to be innovators for other products?

There is considerable evidence showing that early adopters differ from late adopters. Findings from different areas of research are reviewed by Rogers and Stanfield (1968); and several studies of marketing innovators have found similar differences. (Robertson, 1971). However, most studies report weak relationships and on the whole, it seems that no variables apply uniformly to all products.

This observation suggests that innovativeness in one area does not automatically imply innovativeness in other areas, a conclusion in line with the results from the few studies which have directly explored the amount of overlap between innovativeness in different areas (Wärneryd, 1965; Robertson & Myers, 1969; and Arndt, 1968a). They all find practically no overlap among innovativeness in different areas. Therefore, there is little support for innovators as a special market segment.

7. Opinion Leaders

It is a common assumption in communication research that some people act primarily as "opinion leaders" and others primarily as receivers ("followers"). As with innovators the influentials might constitute a special market segment composed of consumers with special values and perceptions.

Findings suggest, however, that opinion leadership does not correlate closely with socioeconomic and demographic variables. Illustrative findings are reported by Myers and Robertson (1969). In an extensive study of 12 products, they found only small correlations with demographic variables and no

single variable was significant for all products, and with regard to personality variables even fewer relationships have been found. (King and Summers, 1969; and Myers and Robertson, 1969).

Several researchers have studied the amount of overlap among opinion leaders in different product areas and found only little overlap between unrelated products and only modest overlap among related products.

The concept of opinion leaders rests upon the assumption that some consumers primarily act as sources of personal communication whereas others primarily are receivers. However, few studies have been concerned with the extent to which the opinion leadership studied results in communication between "leaders" and "followers," and recent studies have suggested that people who predominantly act as receivers ("the followers") are rare (Wärneryd, 1965; Cerha, 1967; and King & Summers, 1969).

These authors suggest that in most product areas from 60% to 80% of all consumers can be characterized either as both frequent receivers and initiators of communication or as infrequent receivers and initiators of communication. When this is taken together with some of the more consistent characteristics of opinion leaders; their interest in innovations, their better knowledge, and their more frequent exposure to mass communication, it appears that rather than distinguishing among opinion leaders and followers, one should distinguish among consumers engaging in more or less personal communications about the product. Basically those who are interested in a given product also are those who talk about it, and in the process of doing this they provide information for others as well as they acquire additional information for themselves.

8. Purchase Characteristics

Several authors have studied relationships among different aspects of the same consumer's behavior. For example, there is some evidence that brand loyalty is positively related to the market share of the brand (Schuchman, 1968), that "deal-proneness" is related to the number of different brands purchased; to the number of units purchased; and to the brand loyalty (Webster, 1965). Similarly Kollat and Willet (1967) found impulse purchases to be related to number of items purchased to the number of members in the shopping party, and to variables reflecting the structure of the transaction (major or minor purchasing trip); Frank et al. (1964) showed relationships between innovative behavior and purchase characteristics; and Rao (1969b) found brand loyalty and private brand proneness to be related to store loyalty.

Relationships of this kind may be useful in some attempts to identify special market segments, but it is usually a problem that consumers with special purchase patterns can rarely be identified in other ways than through their purchase behavior. For example, it is of limited use to know that heavy users tend to be more loyal than light users if neither loyal nor heavy users can be identified. But it is a common observation that neither loyal, deal-prone, private brand loyal, heavy users nor innovative consumers, etc. are easily identified. (Frank, 1968).

9. Image Congruence Theories

Consumers' perceptions of purchase alternatives are reflected in the cognitive relationships between on the one hand brands, products, stores, etc. and on the other hand aroused values. Similarly, the consumer's perception of himself can be described in terms of the perceived relationships between the concept

of the self and valued concepts. Several authors have suggested that the consumer selects products which are perceived as congruent with his self-image. To the extent that such a relationship can be proven, the self-image would be a valuable set of variables to work with in segmentation studies.

A number of studies have attempted to validate the image congruence hypotheses. Most of these have tested one or both of the following two propositions: (1) there are significant differences in the way in which products are perceived, (2) those products which the consumer owns or prefers have images which deviate less from his self-image than the images of the products which he does not own or does not prefer. Several studies have dealt with automobile brands and supportive evidence is reported by Jacobsen and Kossoff (1963), Birdwell (1964), Grubb and Hupp (1968), and Ito (1967).

To test whether knowledge of the consumers' perception of himself together with information about his images of brands makes it possible to predict his choices, both images should be measured before the choice. This has never been tried. What comes closest is the study reported by Ito (1967). In a nationwide probability sample of car owners, 577 Ford and Chevy owners who were planning to purchase a new Ford or Chevy were identified. Based upon measures of self- and product images, it was possible to predict from 51 to 66 percent of the purchase intentions correctly. These percentages are not very high, but of those who intended to switch brands, from 82 to 96 percent were classified correctly.

Another attempt to prove the significance of the self-image rests upon the following reasoning: Since it is not likely that the consumer will change his perception of himself following purchase decisions, significant differences in the self-images of consumers who have chosen different brands can be expected also to have existed before the brand was chosen. Such differences have been identified by Grubb and Hupp (1968) in a study where they compared Pontiac owners and Volkswagen owners. They found that Pontiac owners rated themselves significantly higher on dimensions which were positively associated with the Pontiac, whereas the Volkswagen owners rated themselves significantly higher on dimensions which were positively associated with the Volkswagen.

Altogether some positive evidence has been reported supporting the existence of a relationship between self-images and images of brands and products purchase, and it is possible that improved measurement techniques may strengthen this further.

10. Backwards Segmentation

Based upon information about consumers consumption and purchase patterns, and upon product perceptions, it has been tried to classify consumers. Large scale factor-analysis has made this possible and the approach is in line with the conclusions from several of the previous sections: attempts with personality inventories have suggested that consumer types should be identified starting in analysis of the behavior of consumers; attempts to make predictions based upon socio-economic and demographic studies of different purchase and consumption variables have shown that many of these are correlated and finally the proposal that significant aspects of consumer's conceptual structures may be reflected in their overall life-styles suggests that more systematic utilization of information about different aspects of consumer behavior may lead to meaningful classifications.

Several studies have been reported. Wilson (1966) used factor analysis to identify 20 different variables reflecting aspects of the respondents' product

perceptions and similar findings are reported by Pessemier and Tigert (1966). In this research 14 interest and 8 personality factors were identified, many of which closely resemble those identified in the Wilson study. Later the same authors (Bass, Pessemier & Tigert, 1969; and also Welb, 1968) have worked with purchase data alone. Here again product oriented factors emerged closely resembling those reported in other studies.

All of these studies suggest that it is possible to characterize consumers meaningfully based upon information about their perceptions and consumption and purchase patterns. In spite of considerable differences in the samples and in the type of data which have been used, many almost identical factors have been identified.

11. Interests and Values

In several ways the research reviewed in the previous pages has pointed at the importance of interests and values factors governing consumer behavior. It was suggested that innovativeness may be seen in relation to interests and that the frequency with which personal and other information sources are attended to, vary with interest in the issue. Finally, image congruence studies and backward segmentation has pointed at factors of the value interest type.

It is not surprising that interest shows up as an important variable when it is realized that interest in an issue implies that the consumer has a number of important and positive values in relation to the topic. It is more surprising that very little research has been concerned directly with this variable.

Only few studies have been reported.

Cerha (1967) obtained interest scores for 91 products on simple seven point scales and found considerable variations among products, as well as close relationships between interest and exposure to mass and personal communication. Moreover, he found relatively small intercorrelations among the different product areas, and based upon factor analysis, he was able to identify seven highly independent interest areas. In another study Pennington and Peterson (1969) used the "Strong Vocational Interest Blank" (Cambell, 1966). Being constructed primarily for purposes of counselling and personal selection, this interest questionnaire would be expected to have the same shortcomings as general personality inventories. Nevertheless, based on the most productive items in the test, and using discriminant analysis, the authors were able to make correct predictions of choices among vacation trips and savings forms in 72 to 80% of the cases. Unfortunately, the products chosen were rather special, but further research with this or similar interest test batteries seems promising. Particularly if tests are developed which are especially relevant to consumer behavior, our ability to predict consumer choices may improve significantly.

Segmentation Variables: Summary

The preceding review suggests that the variables most commonly used in segmentation studies can be grouped along two dimensions. On the one hand, one can distinguish among behavioral and psychological variables. On the other hand it is possible to see the variables as being more or less specific to the ultimate purpose with segmentation studies: to divide consumers into groups behaving differently as consumers. This two way classification is shown in Figure 1.

	Psychological Variables	Behavioral Variables
General variables	General Personality Tests Self image Specific Personality Variables	Social Demographic Variables Social Class Life Cycle Opinion Leadership Innovators
Specific consumer variables	Values, Interests Product Perception	Buying and Consumption Behavior Brand Choices

Figure 1. Possible Segmentation Variables: An Overview.

With regard to the more psychologically oriented variables, the precedent review suggests that the more product specific variables are the most promising.

Among the behavioral variables, the more specific ones have in a number of contexts proven useful. However, practical applications of these variables often raise problems. It is normally difficult to identify marketing strategies aimed at market segments defined in terms of the behavioral variables to be influenced by the marketing strategies chosen. Normally, the researcher tries to define the segments identified in this way in terms of some attitude or interest variables which can be deduced from the behavioral variables grouped together.

For this reason in a number of segmentation studies we have chosen to work with different kinds of product related interest, attitude and perceptual variables. Below some of our experiences are reviewed.

Backwards Segmentation With Interest, Attitude And Perceptual Variables

Backwards segmentation was described above as segmentation based upon those behavioral interest and attitude variables in which one is ultimately interested. In the simple case with only one variable involved, backwards segmentation is trivial. For example, dividing consumers into heavy users and light users of a product based upon a single measure of the amount used of the product, is a simple form of backwards segmentation. However, what is most commonly discussed as backwards segmentation is the case where a number of different variables form the basis for the segmentation, and here the inter-relationships among the variables become critical. As an operational definition of backwards segmentation the following is proposed: Backwards segmentation is the process in which segments are established departing from those variables which are dependent in relation to the project.

A Segmentation of TV-Viewers

In this study the major emphasis is on a comparison between the use of two different techniques with the same set of dates.

In one of a regularly mode TV audience surveys, carried out for the Danish Broadcasting Corporation, special interest was paid to the news programs. Criticism had been raised about the TV-news coverage. Particularly the reporters were said to be difficult to understand, using an academic language and being politically biased. To examine the extent to which this criticism existed among the viewers also, a number of questions about TV's news coverage were included in the spring survey of 1971. The questions were concerned with how viewers regard the coverage of local news, foreign news, objectivity, actuality, the use of filler items, understandability, use of foreign words, unclear speaking, etc. With these items no real support for the criticism was found, neither in the total population nor in any geographical, socio-economic or other segments. For this reason it was decided to use a backwards segmentation technique in an attempt to test if segments existed in which the criticism could be identified.

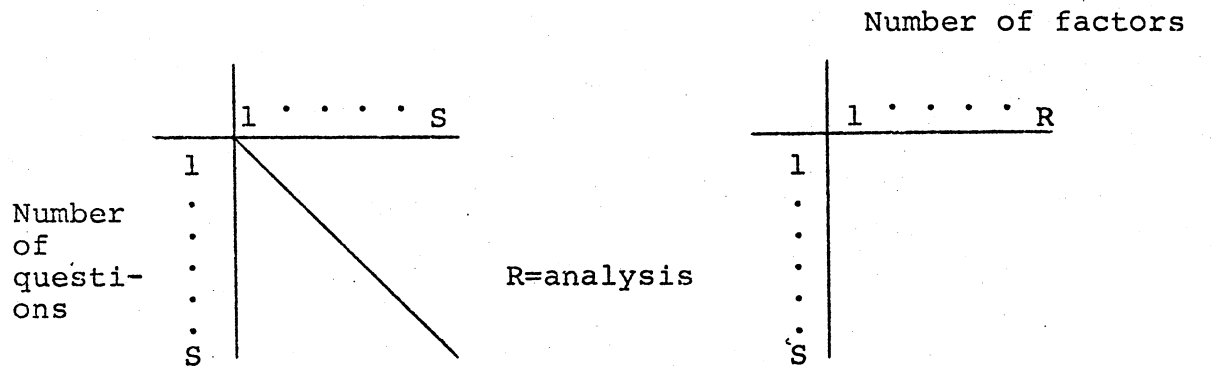
To identify the dimensions along which the 1500 respondents evaluate the newscasts, a normal (R) factor analysis was carried out. In this way three major dimensions were localized. Briefly described they are:

- I Quality of information
- II Understandability
- III Type of information

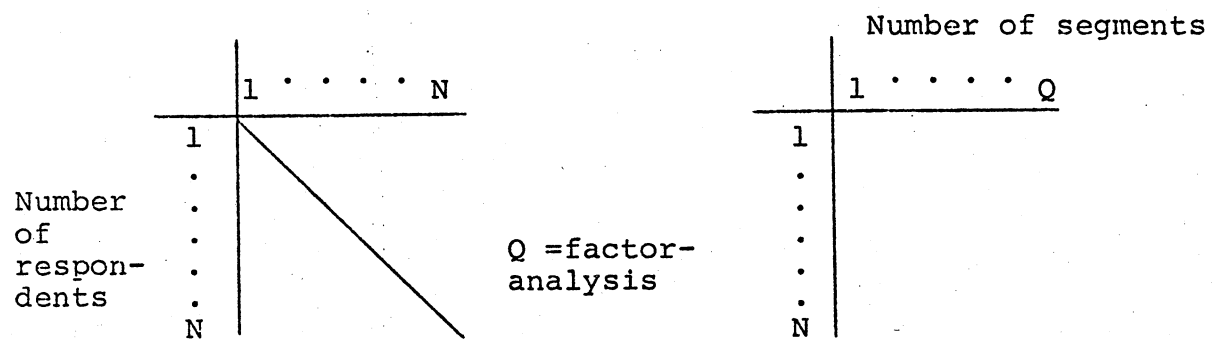
The main results of this analysis are presented in Table 1.

	1.	2.	3.
1. Broad coverage of local news	.826	-	-
2. Broad coverage of international news	.825	-	-
3. Objective	.654	-	-
4. Of current interest	-	-	÷.751
5. Use often filler	-	-	.610
6. Easy to understand	-	÷.699	-
7. Use many foreign words	-	.648	-
8. Speak indistinctly	-	-	.597
9. Trot out one's favourite idea	-	.526	-

Subsequently a Q-factor analysis was carried out in an attempt to identify segments having special viewers along the three major dimensions. Briefly



2a = R. Factor analysis



2b = Q. Factor analysis

Fig. 2. Original and reduced correlation matrices in R and Q factor analysis.

	Total Sample	Segments				
		1.	2.	3.	4.	5.
1. Broad coverage of local news	4.36	5.34	2.70	3.25	4.36	4.00
2. Broad coverage of international news	4.47	5.29	3.25	3.50	4.36	4.92
3. Objective	4.74	5.29	4.85	3.94	5.07	4.67
4. Of current interest	5.27	5.56	5.65	5.41	5.79	4.92
5. Use often filler	3.72	2.98	2.25	5.06	3.57	4.83
6. Easy to understand	5.31	5.74	5.60	5.06	3.21	5.67
7. Use many foreign words	3.78	3.24	3.65	4.76	5.00	2.25
8. Speak indistinctly	1.64	1.33	1.75	1.41	1.93	1.25
9. Trot out One's favourite idea	3.41	2.29	4.74	3.19	4.71	5.45
% of variance explained	100%	43%	13%	11%	10%	8%

TABLE 2

Profile for sample and segments.

A high score means that respondents tried to open with the statements (6 point Likert Scales are used)

the nature of the Q-factor analysis can be understood as follows. Whereas the R analysis departs in a matrix of correlations like the one shown in Figure 2a containing correlations between responses computed over respondents, the Q analysis works with a matrix like the one shown in Figure 2b. Here the correlations are between respondents and computed so that two respondents with the same responses will have a correlation of 1.00 and two respondents with very different responses will have a low correlation.

The computational procedure, however, is the same in Q and R analysis.

In this way 9 segments were identified (Table 2). Of those, however, the 5 largest contained 86% of the respondents and the results are shown in Table 2. Among the segments, the first (including a little less than 50% of all respondents) can be described as satisfied viewers. They are more positive than the average on all dimensions. The remaining four groups are characterized by their negative evaluation along one or more of the dimensions. Group two is claiming that rather than giving a broad coverage of local and international news, biased news are broadcasted. The group constitutes 15% of all respondents. Group 3 is also dissatisfied with the news coverage, but rather than seeing the material broadcasted as biased, they perceive it as "filler items". This segment constitutes 13% of the respondents. The 2nd last segment does not analyse the news coverage as such, but it is claimed that it is biased and less relevant. It is noteworthy that this group to a significant extent says that only few foreign words are used, suggesting that it is a more educated group. Finally the last group (4) says that many foreign words are used, that it is not easy to understand and that it is biased. Seemingly this segment constitutes a group which has difficulties in understanding what is being broadcasted.

It should be noted also, that the segments differing significantly in their view on Radio and TV-news are quite similar in their socio-economic characteristics. The only significant difference lies in the first segment which tends to consist of more males in the Copenhagen area.

In an attempt to test the validity of the grouping mode with the Q analysis a sample of 150 respondents was reanalysed using a different clustering procedure. In this case comparisons among respondents was based upon rank-order correlations between respondents, and the matrix of similarities constructed in this way was then used as input for The Johnson's hierarchical clustering procedure. In this way 5 clusters plus 16 unclassifiable respondents were identified. Particularly the 5 clusters agree very closely with the 5 major segments identified in the Q analysis. There were only few and insignificant differences in the average scores and the overlap between groups runs as high as 60-90%. That is, of the respondents assigned to one segment by the hierarchical clustering procedure from 60 to 90% were assigned to the the corresponding group in the Q analysis.

A Comparison of Segments Established Based Upon Different Types of Data

To understand the segmentation problem encountered in this study it is necessary to say a few words about some of the results from another part of the study.

The second study to be reported deals with educational toys. In a given market, demand for the product is highly dependent upon two things. First, the number of consumers adopting the product and secondly the frequency with

SPECIFIC IMAGE SEGMENTS

ATTITUDE SEGMENTS	Seg- ment No.	1	2	3	4	5	6	Total
	1	59	8	8	22	0	3	100%
	2	39	48	9	0	4	0	100%
	3	35	13	13	17	22	0	100%
	4	65	8	27	0	4	0	100%
	5	26	19	14	5	22	14	100%

(N = 200)

Table 3

Percentage of respondents from segments established based upon general attitudes falling in respondents established based upon specific image variables.

which those who adopt continues to purchase it. The product sales differ significantly between England and Germany. To examine the extent to which this was ascribable to differences in the market structure a comparative study was designed. In this study measurements were obtained on 7 point likert scale. The questionnaire included 35 general attitude statements relating to child rearing, purchases of toys, etc. Moreover, 43 specific image variables revealing the perception of the product were included. With these sets of variables using R-factor analysis it was found that the structure in the general attitudes as well as in the specific images were quite similar in the two countries.

Moreover, in the segments subsequently identified separately in England and Germany the same conceptual structures applied.

Actually the factors were so alike that it was a natural way of carrying out the comparisons between segments to run a joint factor analysis on the combined data from the two countries, resulting in one common image structure (14 factors) and one common general attitude structure (12 factors). Further comparisons could be made with the use of these common factors.

With regard to the general attitudes compared with the specific image it was found that significant differences were found between the two countries in 6 and 8 factors. However, when the variables were related to the amount of the toy in the household (by means of stepwise regression analysis) and to propensities to buy more of the toy, only the specific image variables had any importance. Opposite, when discriminant analysis was used to analyse ownership versus no ownership, the most important variables turned out to be general attitude ones. Seemingly, are the two sets of variables playing a different rate and the question relating to further segmentation was, what set of variables to use.

With the results from the previous study in mind (and those from a similar one on automobiles) it was decided to use only Q-factor analysis, and this was carried out on both sets of data separately. Here only results from one of the countries are shown. With the image data 5 segments were identified and with the general attitude data 6 segments resulted. How the segments compare can be seen from Table 3, where it appears that even though some agreement exists (the zero-hypotheses that the two groups of segments are unrelated is rejected with $p \geq 0.999 - X^2$ test) there is no close correspondance between the segments. Closest come the two segments, number two having approximately 50% of the respondents in common. In further analysis both sets of the segments were studied in relation to amount of the toy at home and propensity to purchase. With regard to the first variable only small differences were found among the segments. This applies to both types of segments. With regard to propensities to buy, however, there were large and significant differences between the segments derived from the image data. This is shown in Table 4 where the four most important segments are shown.

The first segment can be described as overall possible, the second as overall negative and the third and the fourth as partially negative. Finally, it should be noted that there are no significant differences between the segment on socio-economic criterion.

The implications for the company of these findings shall not be discussed here, but some conclusions concerning the technique of backwards segmentation should be provided.

Segment	Total	1	2	3	4
Size of segment	100%	44%	19%	15%	10%
Propensity to buy 1 = sure -6 = sure not to	2.5	1.8	3.0	2.6	2.6
Target group understanding	2.99	3.70	1.59	2.19	4.68
Satisfaction with use	3.17	3.67	2.05	4.06	1.90
Price factor	4.55	5.19	3.68	3.06	3.00
Quality	3.06	3.67	2.55	4.19	2.30

Table 4. Four major segments and their perception of the product.

Conclusions

Two basic questions facing the researcher when dealing with backwards segmentation relate to the choice of the data handling technique and the choice of variables to work with. The findings reported here suggest that the latter question is far more crucial than the first. Whereas choice of technique to a large extent may be a question of the type of data available (metric/nonmetric, etc.) the choice of variables to work with turns out to be highly critical. For that reason it may be advisable to put more time into the latter question.

Another conclusion established in the studies reported here is that backwards segmentation is useful in the sense that it makes it possible to identify segments, which differ in relation to the product studied. Moreover in the two studies reported here (and a couple of others not described) it seems that the segments which emerge normally will consist of one totally positive and a number of segments being negative for different reasons.

Finally, a somewhat negative--but not unexpected--conclusion can be mentioned. Seemingly the segments thus identified do not differ largely in terms of more traditional variables such as income, age, sex, etc.

Footnotes

1. T. Bak-Jensen A/S/A.I.M.-Marketing is a subsidiary of Booz, Allen and Hamilton Inc.
2. The variable should not be confused with 'specific self-confidence' which reflects the individual's confidence in his ability to cope with a specific problem with which he is faced (for a discussion see Cox and Bauer, 1964).

References

- Advertising Research Foundation. Are There Consumer Types? New York: Advertising Research Foundation, 1964.
- Aliport, G. W., Vernong, P. E., Lindzey, G. Study of Values. Boston, Mass.: Houghton-Mifflin Company, 1960.
- Arndt, J. A Study of Work of Mouth Advertising. Markedskommunikation. 1967, 4, 94-117.
- Arndt, J. New Product Diffusion: The Interplay of Innovativeness, Opinion

- Leadership, Learning, Perceived Risk and Product Characteristics. *Markeds-kommunikation*, 1968a, 5, pp. 1-9.
- Arndt, J. Profiling Consumer Innovators. In J. Arndt (Ed.), *Insights Into Consumer Behavior*. Boston: Allyn and Bacon, 1968b.
- Birdwell, E. A. Influence of Image Congruence on Consumer Choice. L. G. Smith (ed.): *Reflections on Progress in Marketing*. Chicago: American Marketing Association, 1964.
- Boulding, E. Orientation Toward Achievement or Security in Relation to Consumer Behavior. *Human Relations*, 1960, 13, 365-82.
- Brim, O. G.; Glass, D. G.; Lavin, D. E.; & Goodman, N. *Personality and Decision Processes*. Stanford, Calif.: Stanford University Press, 1962.
- Brim, O. G.; Hoff, D. B. Individual and Situational Differences in Desire for Certainty. *Journal of Abnormal and Social Psychology*, 1957, 54, 225-29.
- Brody, R. P.; Cunningham, S. M. Personality Variables and the Consumer Decision Process. *Journal of Marketing Research*, 1968, 5, 50-57.
- Cambell, D. P. *Manual for the Strong Vocation Interest Blank*. Palo Alto, Calif.: Stanford University Press, 1966.
- Cerha, J. *Selective Mass Communication*. Stockholm: P. A. Norstedt and Soner, 1967.
- Claycamp, J. H. Characteristics of Thrift Deposit Owners. *Journal of Marketing Research*, 1965, 2, 163-70.
- Cox, D. F. The Influence of Cognitive Needs and Styles on Information Handling in Making Product Evaluations. In D. F. Cox (Ed.) *Risk Taking and Information Handling in Consumer Behavior*. Boston: Grad. School of Business Administration, Harvard University, 1967, pp. 370-73.
- Cox, D. F.; Bauer, R. A. Self Confidence and Persuadability in Women. *Public Opinion Quarterly*, 1964, 28, 453-66.
- Cunningham, S. M. Perceived Risk as a Factor in Product-Oriented Word-of-Mouth Behavior: A First Step. In L. G. Smith (Ed.), *Reflections on Progress in Marketing*. Chicago: American Marketing Association, 1964.
- Cunningham, S. M. Perceived Risk and Brand Loyalty. In D. F. Cox (Ed.), *Risk Taking and Information Handling in Consumer Behavior*. Boston: Graduate School of Business Administration, Harvard University Press, 1967a.
- Cunningham, S. M. The Major Dimensions of Perceived Risk. In D. F. Cox (Ed.), *Risk Taking and Information Handling in Consumer Behavior*. Boston: Graduate School of Business Administration, Harvard University, 1967b.
- Donnelly, Jr., J. H. Social Character and Acceptance of New Products. *Journal of Marketing Research*, 1970, 7, 111-13.
- Edwards, A. L. *Personal Preference Schedule: Manual*. New York: The Psychological Corporation, 1954.
- Evans, F. B. Psychological and Objective Factors in the Prediction of Brand Choice. Ford versus Chevrolet. *Journal of Business*, 1959, 32, 340-69.
- Evans, F. B. Ford versus Chevrolet: Park Forest Revisited. *Journal of Business*, 1968, 41, 445-59.
- Frank, R. E. Is Brand Loyalty a Useful Basis for Market Segmentation? *Journal of Advertising Research*, 1967a, 7, 27-33.
- Frank, R. E. Correlates of Buying Behavior for Grocery Products. *Journal of Marketing*, 1967b, 31, 48-53.
- Frank, R. E. Market Segmentation Research: Findings and Implications. In F. M. Bass, C. W. King and E. A. Pessemier (Eds.), *Applications of the Sciences in Marketing Management*. New York: John Wiley and Sons, Inc., 1968.
- Frank, R. E. & Boyd, H. W. Are Private Brand Prone Grocery Customers Really Different? *Journal of Advertising Research*, 1965, 5, 27-35.
- Frank, R. E.; Green, R. E.; Sieber, H. F. Household Correlates of Purchase Price for Grocery Products. *Journal of Marketing Research*, 1967, 4, 54-58.

- Frank, R. E., Massy, W. F., & Morrison, D. G. The Determinants of Innovative Behavior with Respect to a Branded, Frequently Purchased Food Product. In G. Smith (Ed.), *Reflections on Progress in Marketing*. Chicago: American Marketing Association, 1964, 312-23.
- Graham, S. Class and Conservation in the Adoption of Innovation. *Human Relations*, 1956, 9, 91-100.
- Grubb, E. L., & Hupp, G. Perception of Self, Generalized Stereotypes and Brand Selection. *Journal of Marketing Research*, 1968, 5, 58-63.
- Hovland, C. I., & Janis, I. L., (Eds.), *Personality and Persuadability*. New Haven: Yale University Press, 1959.
- Ito, Pikuma. Differential Attitude of New Car Buyers. *Journal of Advertising Research*, 1967, 38-42.
- Jacobsen, E., & Kossoff, J. Self Percept and Consumer Attitudes Toward Small Cars. *Journal of Applied Psychology*, 1963, 47, 242-45.
- Kassarjian, H. H. Social Character and Differential Preference for Mass Communication. *Journal of Marketing Research*, 1965, 2, 146-53.
- Kassarjian, W. M. A Study of Riesman's Theory of Social Character. *Sociometry*, 1962, 25, 213-30.
- Kernan, J. B. Choice Criteria, Decision Behavior, and Personality. *Journal of Marketing Research*, 1968, 5, 155-64.
- King, C. W., & Summers, J. O. Technology, Innovation and Consumer Decision Making. In R. Moyer (Ed.), *Changing Marketing Systems*. Chicago: American Marketing Association, 1967, 63-68.
- Kollat, D. T., & Willett, R. T. Customer Impulse Purchase Behavior. *Journal of Marketing Research*. 1967, 4, 21-31.
- Koponen, A. Personality Characteristics of Purchasers. *Journal of American Research*, 1960, 1, 6-12.
- Kuehn, A. A. Demonstration of a Relationship Between Psychological Factors and Brand Choice. *Journal of Business*, 1963, 36, 237-41.
- Kuehn, A. A. Mathematical Models of Consumer Behavior. In J. W. Newman (Ed.), *On Knowing the Consumer*. New York: John Wiley and Sons, Inc. 1966.
- Marcus, A. S. Obtaining Group Measures from Personality Test Scores: Auto Brand Choice Predicted From the Edwards Personal Preference Schedule. *Psychological Reports*, 1965, 17, 523-31.
- Martineau, P. D. Social Class and Spending Behavior. *Journal of Marketing*, 1958, 121-30.
- Martineau, P. D. Customers Shopping Center Habits Change Retailing. Editor and Publisher, October 26, 1963, p. 11.
- Massy, W. F., Frank, R. E., & Lodahl, T. *Purchasing Behavior and Personal Attributes*. Philadelphia: University of Pennsylvania Press, 1968.
- Morgan, J. N. The Achievement Motive and Economic Behavior. In J. W. Atkinson and N. T. Feather (Eds.), *A Theory of Achievement Motivation*. New York: John Wiley and Sons, Inc. 1966.
- Murray, H. A. *Explorations in Personality*. New York: Oxford University Press, 1938.
- Myers, J. G. Determinants of Private Brand Attitudes. *Journal of Marketing Research*, 1967, 4, 73-81.
- Ostlund, L. E. The Role of Product Perceptions in Innovative Behavior. Paper presented at Fall Conference of American Marketing Association, Cincinnati, Ohio, 1969.
- Pennington, A. L., & Peterson, R. A. Interest Patterns and Product Preferences: An Exploratory Analysis. *Journal of Marketing Research*, 1969, 6, 284-90.
- Phares, E. J., & Davis, W. L. Breadth of Categorization and the Generalization of Expectancies. *Journal of Personality and Social Psychology*, 1966, 4, 461-63.
- Popielarz, B. T. An Exploration of Perceived Risk and Willingness to Try a New Product. *Journal of Marketing Research*, 1967, 4, 368-73.

- Rao, T. R. Are Some Consumers More Prone to Purchase Private Brands? *Journal of Marketing Research*, 1969, 6, 447-50.
- Rich, S. U., & Jain, S. L. Social Class and Life Cycle Predictions of Shopping Behavior. *Journal of Marketing Research*, 1968, 5, 41-49.
- Riesman, D. *The Lonely Crowd*. New Haven: Yale University Press, 1961.
- Rizzo, J. R., & Naylor, J. C. The Factorial Structure of Selected Consumer Choice Parameters and Their Relationship to Personal Values. *Journal of Applied Psychology*, 1964, 48, 241-48.
- Robertson, T. R., & Myers, J. G. Personality Correlates of Opinion Leadership and Innovative Buying Behavior. *Journal of Marketing Research*, 1969, 6, 164-68.
- Rogers, E. M., & Stanfiels, J. D. Adoption and Diffusion of New Products: Emerging Generalizations and Hypotheses. In F. M. Bass, E. A. Pessemier and C. W. King (Eds.), *Applications of the Sciences in Marketing Management*. New York: John Wiley and Sons, Inc., 1968.
- Rotzoll, K. B. The Effect of Social Stratification on Market Behavior. *Journal of Advertising Research*, 1967, 7, 22-27.
- Ruch, D. M. Limitations of Current Approaches to Understanding Brand Buying Behavior. In J. W. Newman (Ed.), *On Knowing the Consumer*. New York: John Wiley and Sons, Inc., 1966.
- Schuchman, A. Are There Laws of Consumer Behavior? *Journal of Advertising Research*, 1968, 8, 19-28.
- Schuchman, A., & Perry, M. Self Confidence and Persuadability in Marketing. *Journal of Marketing Research*, 1969, 6, 146-55.
- Scodel, A., Ratoosh, P., & Minas, J. S. Some Personality Correlates of Decision Making Under Conditions of Risk. *Behavioral Science*, 1959, 4, 19-28.
- Scott, E. M. Personality and Movie Preferences. *Psychological Reports*, 1957, 3, 17-18.
- Steiner, G. A. Consumer Behavior: Where Do We Stand? A Psychologist's Appraisal. In J. W. Newman (Ed.), *On Knowing the Consumer*. New York: John Wiley and Sons, Inc. 1966.
- Thurstone, L. L. *Examiner Manual for the Thurstone Temperament Schedule*. Chicago: Science Research Associates, 1953.
- Tucker, W. T., & Painter, J. J. Personality and Product Use. *Journal of Applied Psychology*, 1961, 45, 325-29.
- Warner, W. L., & Lunt, P. S. *The Social Life of a Modern Community*. New Haven: Yale University Press, Yankee City Series, Vol. 1, 1941.
- Wärneryd, B. *Innovation, Indflydelse och Information (with an English summary)*. Stockholm, Sweden: Almqvist och Wicksell, 1965.
- Webster, F. E., Jr. The Deal-Prone Consumer. *Journal of Marketing Research*, 1965, 2, 186-89.
- Wells, W. D. Children as Consumers. In J. W. Newman (Ed.), *On Knowing the Consumer*. New York: John Wiley and Sons, Inc., 1966.
- Wells, W. D. Backward Segmentation. In J. Arndt (Ed.), *Insights Into Consumer Behavior*. Boston: Allyn and Bacon, 1968.
- Westfall, R. L. Psychological Factors in Predicting Product Choice. *Journal of Marketing*, 1962, 26, 34-40.
- Wicks, J. H., & Nelson, C. C. Preliminary Investigation of Some Psychological Determinants of Consumption Propensity. *Southern Economic Journal*, January, 1967, 383-87.
- Wilding, T., & Bauer, R. A. Consumer Goals and Reaction to a Communication Source. *Journal of Marketing Research*, 1968, 5, 73-77.
- Wilson, L. L. Homemaker Living Patterns and Marketplace Behavior--a Psychometric Approach. In J. S. Wright and J. L. Goldstucker (Eds.), *New Ideas for Successful Marketing*. Chicago: American Marketing Association, 1966.

DISCRIMINATING BETWEEN STOCHASTIC MODELS OF BRAND CHOICE:
MINIMUM CHI-SQUARE AND BAYESIAN METHODS¹

Robert Blattberg and Subrata Sen
Graduate School of Business
University of Chicago

During the last ten years, several stochastic models of consumer brand choice have been proposed by various researchers. These models range from the Bernoulli, Markov, Linear Learning, and Probability Diffusion models described in Massy, Montgomery, & Morrison (1970), to the more recently developed New Trier model (Aaker, 1971) and Dual Effects model (Jones, 1971). One of the difficult problems in model development is measuring how well a particular model "fits" a set of empirical data (e.g., consumer panel data). The underlying model assumptions often do not clearly indicate which model is appropriate in a given situation. Hence, it is important to develop systematic techniques for discriminating between alternative models of brand choice.

During the last few years, the most commonly used measure of model fit in marketing has been the minimum chi-square statistic. This statistic was first used in marketing by Morrison (1966) to estimate the parameters of a stochastic brand choice model. The model parameters are estimated by minimizing this chi-square statistic. The minimum chi-square statistic can then be used as a measure of model fit. Recently, however, researchers have proposed alternative criteria for model fit. For example, Aaker (1970) has suggested that model tests should not be restricted to the usual goodness-of-fit test but that each model should be judged in terms of how well it can predict market shares in future time periods.

This paper will describe two techniques of model discrimination:

(1) Rao's chi-square technique, and (2) a Bayesian technique. Two-state heterogeneous Bernoulli and Markov models will be used to illustrate the discriminatory power of these two techniques. The two-state heterogeneous Bernoulli model is merely a constrained version of the two-state Markov model. If one of the parameters of the Markov model is constrained, the Bernoulli model is obtained. This makes it possible to use Rao's chi-square test to discriminate between these two models (Rao, 1961). The Bayesian discriminatory procedure consists of a Jeffreys test (Jeffreys, 1948). The test results will consist of the posterior odds of obtaining an observed set of consumer purchase data, given the two models.

Initially, the two model discrimination techniques will be tested on artificial data (simulating consumer panel data) generated from a Bernoulli process and a Markov process. Since the model generating the data is known; the success of the two discrimination techniques can easily be determined. The experimentation will be conducted over a wide range of parameters of the Bernoulli and Markov processes. This paper will only report the results obtained using the artificial data. Application of the two discrimination techniques to actual consumer panel data consists of the next phase of the study and will be reported later.

Literature Survey

This section describes the chi-square goodness-of-fit test developed by Morrison (1966) as well as its amended version used by other researchers

like Aaker (1970) and Jones (1970). This description is followed by a discussion of other criteria proposed to evaluate stochastic brand choice models. Aaker (1970) focuses on the importance of a model's ability to predict future market shares while Jones (1970) stresses the importance of a model's ability to provide important insights into the actual mechanisms occurring in the market place.

The Chi-Square Goodness-of-Fit Test

Consumer panel data typically consists of the purchase records over time of a representative sample of N families for a wide variety of product classes. For a given product class, each family's purchase record can be represented by a string of 1's and 0's. A "1" represents purchase of the brand under consideration (henceforth denoted as Brand 1), while "0" represents purchase of any other brand.

Assume that data are available for five consecutive purchases for each of the N families in the panel. Consider the first four purchases. The N families in the panel can be segmented by these four purchases into 2^4 or 16 mutually exclusive and exhaustive categories (e.g., 0000, 0001, 0010, ..., 1011, 0111, 1111). If a five-purchase sequence is used, 2^5 or 32 such categories would be obtained. The panel data can now be represented in terms of N_i (the number of families whose pattern of purchases is represented by category i) and R_i (the number of families with past history i which purchased Brand 1 on the fifth or most recent purchase. For a four-purchase sequence, i ranges from 1 to 16. In general, i ranges from 1 to 2^m where m is the purchase sequence length. The expected number of families (for a particular stochastic brand choice model) with past history i which purchase Brand 1 on the fifth trial is denoted by E_i , where:

$$E_i = N_i P(1|i).$$

$P(1|i)$ is the theoretical probability that a family with past history i will purchase Brand 1 on the next trial for a particular model. $P(1|i)$ is called the conditional model probability and has to be computed for each model.

Morrison's goodness-of-fit statistic is:

$$\chi^2_M = \sum_{i=1}^{2^m} \frac{(R_i - E_i)^2}{E_i} + \sum_{i=1}^{2^m} \frac{(R'_i - E'_i)^2}{E'_i},$$

where R'_i = the number of families with purchase history i who did not purchase Brand 1 on the $(m+1)$ purchase = $(N_i - R_i)$,
 $E'_i = (N_i - E_i)$ and is similarly defined.

The statistic χ^2_M is asymptotically distributed chi-square with 2^m degrees of freedom. Usually, the model's parameters are estimated from the data by minimizing this χ^2_M statistic. If q parameters are estimated from the data, the degrees of freedom are reduced to $(2^m - q)$.

Aaker (1970) and Jones (1970), along with a few other researchers, use a slightly different chi-square statistic. This statistic is defined as follows:

$$\chi^2_A = \sum_{i=1}^{2^m} \frac{(N_i - NV_i)^2}{NV_i}$$

where $N =$ total number of families in the panel $= \sum_{i=1}^{2^m} N_i,$

$V_i =$ probability of purchase sequence i for a particular stochastic brand choice model.

χ_A^2 is asymptotically distributed as chi-square with $(2^m - 1)$ degrees of freedom which are reduced to $(2^m - 1 - q)$ when q model parameters are estimated from the data.

The numerical value of the minimum chi-square statistic cannot be used to compare models with different numbers of parameters. To correct for the different degrees of freedom, one must use the p-level associated with a chi-square statistic. The p-level is defined as:

$$\text{p-level} = \int_x^{\infty} f(x) dx$$

where $f(x)$ is the chi-square distribution with the appropriate degrees of freedom.

A low p-level indicates that the particular model is not a viable representation of the process. However, if the sample size is sufficiently large, only a "perfect" model will avoid a low p-level. Thus, the relative sizes of the p-levels should be used in evaluating the fit of several alternative models to a set of consumer panel data. This is the approach taken by Aaker (1970) in comparing the Linear Learning model with the New Trier model. Jones (1970) uses a similar approach to compare three versions of the Dual Effects model with the Probability Diffusion model and the Linear Learning model.

Alternative Approaches to Model Evaluation

An important objective for most stochastic brand choice models is the prediction of future market shares. However, the chi-square goodness-of-fit test involves only the purchases used to estimate the model parameters. Hence, a model could provide a good fit to the first few purchases used to estimate the model's parameters and yet be quite inaccurate in predicting future market shares. Aaker (1970) provides an interesting example of this possibility. In terms of the chi-square goodness-of-fit test, the Linear Learning model appeared to be superior to the New Trier model for two sets of data from a frequently purchased consumer good product. However, the p-levels for the New Trier model (0.52 and 0.88 compared to 0.54 and 0.99 for the Linear Learning model) were quite respectable. Thus, the New Trier model could not be rejected out of hand. Market share predictions told a different story. After the first four purchases (which were used to estimate its parameters), the Linear Learning model was consistently inferior to the New Trier model in predicting the future market shares of the two brands. A model (like the Linear Learning model) which contains several parameters can often be fitted quite well to a given set of data. However, a true test of a model's viability is its performance over a hold-out sample which was not used to estimate its parameters. The Bayesian technique described in this paper and Aaker's method are both designed to evaluate a model in terms of its predictive ability.

Jones (1970) used panel data on Crest toothpaste to evaluate several models of brand choice. The data covered the period following the August 1960 endorsement of Crest toothpaste by the American Dental Association (ADA). Jones found that the Linear Learning model and the Probability Diffusion model provided better fits to the data compared to the three

versions of the Dual Effects model. Though the p-values of the Dual Effects models were not low enough to discard them immediately, the chi-square goodness-of-fit test did portray them in an unfavorable light. However, a comparative analysis of the parameters of the different models showed that the Dual Effects models were able to indicate complex consumer behavior patterns which could not be adequately modeled by either the Linear Learning model or the Probability Diffusion model. Essentially, the Dual Effects models were able to separate the effect of the ADA endorsement of Crest from the effect of the feedback obtained from purchasing and using the brand. Thus, a model which can provide a better understanding of complex consumer behavior should not be discarded merely because it does not fare as well as alternative models in a goodness-of-fit test.

The next sections describe two other model discrimination techniques: (1) Rao's chi-square technique, and (2) a Bayesian technique. Rao's technique facilitates discrimination between two models, one of which is a constrained version of the other. The Bayesian technique is in the spirit of Aaker's suggestion of judging a model in terms of its power to predict future market shares. The Bayesian technique uses half the data to estimate parameters and uses the remaining data to predict posterior odds of the two models.

Markov and Bernoulli Models

Before discussing the two discrimination techniques, brief descriptions of the Markov and Bernoulli models are provided below. These models are used to illustrate the discriminatory power of the two techniques.

Markov Model of Consumer Behavior

The first-order Markov model states that the last purchase, and only the last purchase, influences the family's current purchase decision. As before, each purchase occasion is represented by a "1" if Brand 1 (as defined earlier) is purchased and by a "0" if some other brand is purchased. This implies that the model of consumer behavior is a two-state, first-order Markov process. If the n^{th} purchase occasion results in the purchase of Brand 1, the probability of purchasing Brand 1 on the next purchase is p_1 while the probability of purchasing some other brand is $(1 - p_1)$. Similarly, if the n^{th} purchase occasion does not result in the purchase of Brand 1, the probability of purchasing Brand 1 on the next purchase is p_2 while the probability of purchasing some other brand is $(1 - p_2)$. These probabilities are known as transition probabilities and can be written in terms of a transition probability matrix:

$$P_M = \begin{matrix} & \begin{matrix} 1 & 0 \end{matrix} \\ \begin{matrix} 1 \\ 0 \end{matrix} & \begin{bmatrix} p_1 & (1 - p_1) \\ p_2 & (1 - p_2) \end{bmatrix} \end{matrix}$$

Bernoulli Model

The Bernoulli model states that the current purchase decision is not influenced by any past purchases. For each purchase occasion, the family has a probability, p , of purchasing Brand 1 and a probability, $(1 - p)$, of not purchasing Brand 1. The transition probability matrix is:

$$P_B = \begin{matrix} & & 1 & & 0 \\ & & & & \\ & & & & \\ & & & & \\ & & & & \end{matrix} \begin{bmatrix} p & (1-p) \\ p & (1-p) \end{bmatrix}$$

Clearly, the Bernoulli model is a special case of the two-state, first-order Markov model. If $p_1 = p_2 = p$ in P_M , the Bernoulli model is obtained.

Rao's Chi-Square Model Discrimination Technique

Rao's chi-square model discrimination technique is illustrated in terms of the Brand Loyal Markov model postulated by Morrison (1966) and the Compound Beta Bernoulli model described in Massy *et al.* (1970, pp. 60-68). The Brand Loyal Markov model states that each family follows a first-order 0-1 process with transition matrix:

$$P_{BLM} = \begin{matrix} & & 1 & & 0 \\ & & & & \\ & & & & \\ & & & & \end{matrix} \begin{bmatrix} p & (1-p) \\ kp & (1-kp) \end{bmatrix}$$

Compared to P_M , the transition matrix of the general Markov model, $p_2 = kp$. k is a parameter of the model and lies between 0 and 1. k is the same for each family. On the other hand, p has a beta distribution over the families in the population, i.e.,

$$b(p) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{(\alpha-1)} (1-p)^{(\beta-1)} \quad \text{for } 0 < p < 1$$

$$= 0 \quad \text{otherwise}$$

where $\Gamma(\cdot)$ is the gamma function and $\alpha, \beta > 0$.

The mean and variance of p are:

$$E(p) = \frac{\alpha}{\alpha + \beta}$$

$$\text{Var}(p) = \frac{\alpha\beta}{(\alpha + \beta + 1)(\alpha + \beta)^2}$$

Clearly, if $k = 1$, the Brand Loyal Markov model reduces to the Compound Beta Bernoulli model. Hence, the Compound Beta Bernoulli model is a constrained version (where the parameter k is constrained to be 1) of the Brand Loyal Markov model.

If efficient estimates of the parameters of the two models are obtained, Rao's test (Rao, 1961, pp. 32-33) can be used to discriminate between the two models. The parameters of the two models are estimated in this paper by minimizing Morrison's chi-square statistic, χ_M^2 . Minimum chi-square estimates meet the criteria of efficiency, satisfying the requirements for Rao's test. Rao's test consists of the computation of the following test statistic:

$$\chi^2_{R} = \sum_{i=1}^{2^m} \frac{(E_{iM} - E_{iB})^2}{E_{iB}} + \sum_{i=1}^{2^m} \frac{(E'_{iM} - E'_{iB})^2}{E'_{iB}}$$

where E_{iM} and E'_{iM} refer to the expected frequencies (as defined before) for the Markov model and E_{iB} and E'_{iB} are the expected frequencies for the Bernoulli model. The χ^2_R statistic is distributed as chi-square with $(q - r)$ degrees of freedom. q is equal to the number of parameters of the unconstrained model while r refers to the number of parameters of the constrained model. In this particular situation, $q = 3$ (i.e., α , β , and k) for the Markov model, while $r = 2$ (i.e., α and β) for the Bernoulli model.

'Jeffreys' Test for Discriminating Between Markov and Bernoulli Models

To discriminate between a two-state Markov process and a two-state Bernoulli process, we can use Jeffreys' test (Jeffreys, 1948, Chapter V). Jeffreys' test is a Bayesian technique which is briefly outlined below.

Description of the Jeffreys Test

Panel data provide a history of the purchasing behavior (represented by a series of 1's and 0's as defined earlier) of each family in the panel. Let X represent a particular family's purchasing behavior. H_i represents hypothesis i , $i = 1, 2$. Let H_1 represent the hypothesis that the family's purchasing behavior is generated by a Bernoulli process. H_2 signifies that a Markov process generates the family's purchasing behavior. The vector of model parameters (i.e., the elements of the transition probability matrices, P_M and P_B) is represented by θ .

Let $p(H_i)$ = prior probability that hypothesis i is true

$p(\theta|H_i)$ = prior distribution for the parameter θ , given that H_i is true

$p(X|\theta, H_i)$ = likelihood function, given the parameter values θ and that H_i is true

$p(H_i|X)$ = posterior probability that hypothesis i is true.

When H_i is true, the correct action is defined as a_i and thus our loss, $L(a_i, H_i) = 0$. When H_i is true and we mistakenly take action a_j , we incur a loss denoted by $L(a_j, H_i)$ for $i \neq j$. Thus, we define the following loss table:

		Actions	
		a_1	a_2
Hypotheses	H_1	0	$L(a_2, H_1)$
	H_2	$L(a_1, H_2)$	0

To decide which action to take, we use minimum expected loss as our criterion. After observing the sample, X , we choose a_1 if

$$L(a_2, H_1)p(H_1|X) > L(a_1, H_2)p(H_2|X) \quad (1)$$

We can compute the posterior probability that hypothesis i is true from

$$p(H_i|X) = \frac{p(X|H_i)p(H_i)}{\sum_{i=1} p(X|H_i)p(H_i)} \quad (2)$$

where $p(X|H_i)$ is the predictive probability distribution for X given that H_i is true, and

$$p(X|H_i) = \int_{\theta} p(X|\theta, H_i)p(\theta|H_i)d\theta \quad (3)$$

The decision rule (1) can be rewritten as: choose a_1 if

$$\frac{p(H_1|X)}{p(H_2|X)} > \frac{L(a_1, H_2)}{L(a_2, H_1)} \quad (4)$$

Decision rule (4) implies that we should choose action a_1 if our posterior odds ratio favoring H_1 is greater than the loss ratio. Substituting equation (2) for $p(H_i|X)$ in (4), the above decision rule can be expressed as:

$$\frac{p(X|H_1)}{p(X|H_2)} > \frac{p(H_2)}{p(H_1)} \cdot \frac{L(a_1, H_2)}{L(a_2, H_1)} \quad (5)$$

The terms on the right-hand side of (5) are assumed to be known prior to observing the sample, X . The term on the left-hand side of (5) is the ratio of the predictive probability distributions for X , given that H_i is true.

Application of Jeffreys' Test to Markov vs. Bernoulli Process

Application of Jeffreys' test requires the development of the predictive distributions for the Markov and Bernoulli models. Equation (3) indicates that development of the predictive distributions requires the definition of (a) a prior distribution, $P(\theta|H_i)$, and (b) a likelihood function, $P(X|\theta, H_i)$ for each of the two models. For the Markov model, a matrix beta distribution is used for the prior while the likelihood function is represented by a Whittle distribution. These distributions are explained in the Appendix where it is shown that the resulting predictive distribution for the Markov model is the beta-Whittle distribution. For the Bernoulli model, the prior is represented by a beta distribution while a binomial distribution is used for the likelihood function. This results in a beta-binomial predictive distribution for the Bernoulli model (see the Appendix).

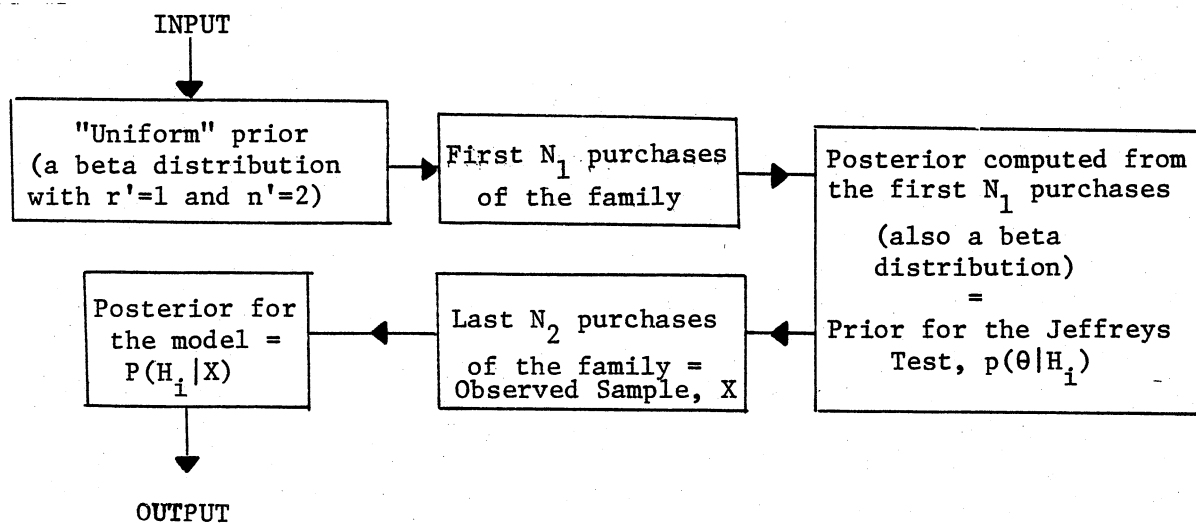
Both prior distributions require knowledge of the parameters of the distributions. For the matrix beta prior used for the Markov process, the parameter matrix, M , must be inputted (M is described in the Appendix). Similarly, the parameters, r' and n' , must be inputted for the beta prior distribution used for the Bernoulli process. Our method for determining these parameters is to split the data into two parts. Assume that we have data on $(N_1 + N_2)$ successive purchases for each family in the panel. The first N_1 purchases are used to determine the parameters of the prior distributions while the remaining N_2 purchases constitute the observed sample, X , which is used for the Jeffreys test.

Natural Conjugate Distribution

To understand how the Jeffreys test is applied, the concept of natural conjugate distributions must be understood. A natural conjugate family of distributions of the parameters θ has the following property: If the prior distribution of θ belongs to a family of natural conjugate distributions, then for any sample n and any values of the observations in the sample, the posterior distribution of θ must also belong to the same family (DeGroot, 1970, p. 159).

To take advantage of this property, we begin with a relatively diffuse prior for each process, observe a sample (the first N_1 purchases of the family), and compute a posterior distribution. The priors we actually use are not diffuse priors but are "uniform" priors. For the Bernoulli process, our prior will be the uniform distribution which is a beta distribution with $r' = 1$ and $n' = 2$, i.e., $f_{\beta}(p|1, 2)$. For the Markov process, we will use an analogous prior, a matrix beta prior with parameter matrix $M' = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. The reason for not using diffuse priors is to avoid zero elements in the parameter matrix of the posteriors. The priors chosen are "close" to diffuse priors and yet avoid the problem of zero elements.

The matrix beta distribution is a natural conjugate prior for \tilde{P} , the transition matrix of a Markov chain. Similarly, the beta distribution is a natural conjugate prior for \tilde{p} , the probability of a success for a Bernoulli process. Hence, the posteriors computed from the first N_1 purchases of the consumer will be of the same family as the priors. Thus, the posteriors computed from the first N_1 purchases will be a matrix beta and a beta distribution for the Markov and Bernoulli models, respectively. These posteriors can now be used as the priors, $P(\theta|H_i)$, for the Jeffreys test which will be run on the last N_2 purchases of the family. The process is illustrated by the following diagram for the Bernoulli model:



Jeffreys' Test Results

Instead of reporting whether we choose H_1 (the Bernoulli model) or H_2 (the Markov model), we will report the posterior odds ratio for the two models which is:

$$\frac{p(H_1|X)}{p(H_2|X)} .$$

This posterior odds ratio for the two models is obtained by multiplying the odds ratios obtained for each family in the panel. The prior odds for the two models will be assumed to be equal, i.e., $p(H_1) = p(H_2) = 1/2$.

A Comparison of Model Discrimination Techniques

This section consists of a brief discussion of the advantages and disadvantages of the Bayesian model discrimination technique described in the previous section. The Bayesian technique is compared first with the chi-square method and then with Aaker's predictive method.

Jeffreys' Test vs. Minimum Chi-Square Method

Compared to the minimum chi-square (MCS) method, the Jeffreys test requires longer purchase histories per family. Applications of the (MCS) method have typically used five purchases per family. In contrast, the Jeffreys test requires 15 to 30 purchases per family. This is necessary because the first part of the purchasing history is used to estimate parameters which are then used to predict the posterior odds ratios of the two models, given the family's last N_2 purchases. However, the information provided by the Jeffreys test is of greater value since it does not merely fit parameters to data (as is done by the MCS method) but also provides an indication of the predictive ability of a model by the use of a hold-out sample. As Aaker (1970) has pointed out, a model which fits the data well (in terms of the MCS method) for the first few purchases may be a poor predictor of future market shares.

Though the MCS method requires only a few purchases per family, the number of families in the sample must be quite large. In most studies, this number has ranged from 500 to over 5,000. This is necessary because the MCS method is a large sample, asymptotic technique. On the other hand, the Jeffreys test requires a sample of only about 50 families.

Because the Jeffreys test requires longer purchase histories, it does not appear to be a useful model discrimination technique for new products. The manager of a new product generally has to determine the underlying model of consumer behavior (and thereby makes predictions about its ultimate success) within a short period of the product's introduction. He usually has to make a go--no-go decision about the product long before he can accumulate purchase histories ranging from 15 to 30. However, for existing products, where long purchase histories are already available, the Jeffreys test appears to be very promising for model discrimination purposes.

Jeffreys' Test vs. Aaker Market Share Prediction Method

The Jeffreys test is very much in the spirit of Aaker's Market Share Prediction method. Both methods use a split-sample approach, using part of the data for parameter estimation and the remaining data for model prediction. However, the Jeffreys test provides a concrete indicator of the

predictive ability of the two models: the posterior odds ratio. Aaker (1970, p. 305) provides graphs of the empirical brand shares and the mean value functions of the models. One must judge the viability of the two models by a visual examination of the graphs. One could devise various indicators of predictive accuracy for Aaker's graphs, e.g., mean square error, mean average deviation, etc. In our opinion, such measures will not provide an indicator as intuitively satisfactory (from the viewpoint of interpretability) as the posterior odds ratio provided by the Jeffreys test.

In summary, the Bayesian Jeffreys test appears to be a useful model discrimination technique for existing products. It does not require a large number of families, and provides an easy-to-understand indicator of a model's predictive ability. Its one apparent shortcoming is that it requires long purchase histories. On the other hand, it appears to be a very general technique. For instance, the application of the Jeffreys test outlined in this paper can handle n -state Markov processes (where $n > 2$) quite easily, whereas Morrison's Brand Loyal Markov model (Morrison, 1966) is essentially limited to two states.

Monte Carlo Simulation Results

The previous section offered a theoretical comparison of the Bayesian and Classical approaches. In this section we study how the two approaches discriminate between the Bernoulli and Markov models.

The data we used are artificially generated Markov data following the heterogeneous Markov process with parameters k , α , and β . When $k = 1$, the data follows a heterogeneous Bernoulli process. For fixed values of α and β the data generated for different values of k are not independent. The same set of uniform random numbers was used to generate the data with numbers in each cell (1111, 1110, ..., etc.) differing only due to different values of k . The reason this was done was to isolate the effect of k on our tests. If, for each value of k , a different set of random numbers was used to generate the data, the results might be attributable to fluctuations in the random numbers rather than changes in k .

For the classical approach, we used Morrison's minimum chi-squared method described earlier to estimate the parameters of the Bernoulli and Markov model. We also computed the chi-squared statistic for Rao's test. The number of observations in our sample is 200.

Table 1 gives the results. We see for $\alpha = 3$, $\beta = 2$ and $k = 1$ the parameters estimates for α and β are slightly upward biased, k is accurately estimated, and the Rao's test would not reject independence at any reasonable significance level. As k decreases our parameter estimates for α and β fluctuate slightly for the Bernoulli model. For the Markov model we see that k is always overestimated. Rao's test never rejects the null hypothesis at reasonable significance levels (the p -value is always less than .73).

The results for $\alpha = 6$ and $\beta = 4$ contrast those of $\alpha = 3$ and $\beta = 2$. For $k = 1$ we see our estimates of α and β are extremely low for the Bernoulli and Markov models compared to the actual values of 6 and 4. The estimate of k is also low, .89, and the p -value for Rao's test is .78. As k decreases from one, the estimates of α and β for both models are still extremely low. The estimates of k are less than the true value for all values of k . We also reject Rao's test of $k = 1$ for all values of k except when $k = 1$.

Table 1
Rao's Test Results

True Values		Bernoulli Estimates			Markov Estimates			Rao's Test
k	α	$\hat{\alpha}$	$\hat{\beta}$	χ^2	$\hat{\alpha}$	$\hat{\beta}$	χ^2	χ^2_R
1.00	3	4.1	3.6	20.33	4.2	3.6	20.20	0.04
.85	3	4.4	4.0	12.67	4.8	4.2	12.68*	0.18
.75	3	4.0	3.9	8.61	5.7	4.6	7.26	1.15
.60	3	2.1	2.7	12.39	2.6	2.9	10.61	1.23
1.0	6	2.3	.9	18.54	2.4	.8	17.93	1.53
.85	6	3.2	1.6	25.05	3.6	1.1	18.21	8.02
.75	6	2.2	1.3	32.72	3.6	1.0	15.39	15.77
.60	6	1.6	1.6	36.04	1.7	.6	15.09	17.71

* The chi-squared value for the Markov model should always be less than the Bernoulli. This value is slightly larger than the Bernoulli chi-squared value because of rounding error.

These above results indicate great ambiguity. In one case, the estimates consistently underestimate k , whereas in the other case, they consistently overstate k . Because for a fixed value of α and β , the same random numbers are used for all k , the results do not offer independent evidence on estimates of k . However, they indicate if there is a bias in k , it persists for all k .

An important result is that Rao's test may not reject $k = 1$ even though k is far from 1. See $\alpha = 3, \beta = 2$. For $k = .6$, the χ^2 statistic is still only 1.23, which has a p-value of .73. Unfortunately, due to the limited number of simulations, we can not give general results about the power of Rao's test.

Turning to the Bayesian procedure, three sets of runs were made. The number of observations used to develop the parameters for the prior is N_1 ; the number of observations to compute the odds ratio is N_2 . The data were generated exactly the same as for the classical approach.

The results are given in Table 2. They indicate for $k = .6$ and $k = 1$ with $N_1 = 15, N_2 = 15$, or $N_1 = 10, N_2 = 10$, the correct model was chosen. However for $k = .85$ and $k = .75$, the Bernoulli model was

Table 2

Jeffreys' Test Results

N_1	N_2	k	α	β	Odds ratio
15	15	1.0	6	4	8.864×10^6
15	15	.85	6	4	1.015×10^5
15	15	.75	6	4	1.036×10^3
15	15	.6	6	4	9.057×10^{-4}
10	10	1.0	6	4	2.388×10^3
10	10	.85	6	4	4.471×10^3
10	10	.75	6	4	3.823×10^4
10	10	.6	6	4	6.740×10^{-2}
10	5	1.0	6	4	1.356×10^3
10	5	.85	6	4	3.198×10^3
10	5	.6	6	4	6.583×10^4

N_1 = sample size to determine prior.

N_2 = sample size to compute odds.

Odds ratio = $\frac{\text{Bernoulli}}{\text{Markov}}$.

chosen when the true model was the Markov. As the sample size for N_1 and N_2 increases from 10, 10 to 15, 15, this result persists. However, as the sample sizes for N_1 and N_2 become large, it can be shown that this bias will disappear. The reason it exists for small samples is that in any cell the frequency count is small for the Markov model. To discriminate between a Markov and Bernoulli model we need to observe a difference between the (1,1) cell and the (2,1) cell as well as the (1,2) and the (2,2) cells. With only 15 observations, we can never observe a large discrepancy. Thus, k has to differ from 1 by enough to make the expected difference between cells (1,1) and (2,1) and between (1,2) and (2,2) large enough to differentiate between the two models.

Summarizing the results, the classical approach for 200 family histories may lead to inaccurate estimates of k , the parameter identifying whether the data come from a Markov or Bernoulli process. For one of the two samples studied, the Bernoulli process is not rejected for any values of k studied. However, the number of replications used in the simulation is so small that it is impossible to generalize these conclusions.

The Bayesian approach discriminates accurately when $k = .6$. For larger values of k (.75, .85) and $N_1 = 15$, $N_2 = 15$, the Bayesian procedure inaccurately gives the posterior odds in favor of the Bernoulli model. These results are due to the small sample sizes chosen. For panel data, it is unreasonable to expect more than 30 observations per family, however. Thus, when using actual marketing data, the Bayesian approach may lead to incorrect classification if k is greater than .75 but less than one. However, when k is sufficiently different from one, the Bayesian approach correctly predicts the true model.

Footnote

1. Research financed in part by National Science Foundation Grant GS-2347. The authors wish to thank Professor Arnold Zellner for his helpful comments.

Appendix

This Appendix describes the prior distributions and the likelihood functions of the Markov and Bernoulli models. It also derives the predictive distributions for the two models.

The Predictive Distribution for a Markov Process: The Beta-Whittle Distribution

This section is based upon Martin (1967, pp. 118-155). From equation (3), the formula for the predictive distribution is:

$$p(X|H_i) = \int_{\theta} p(X|\theta, H_i) p(\theta|H_i) d\theta .$$

For a Markov process, the prior, $p(\theta|H_2)$, will be the matrix beta distribution, while the likelihood function, $p(X|\theta, H_2)$, will be the Whittle distribution. The parameters, θ , are the transition probabilities which constitute the transition matrix, P_M . The sample observations, X , are observed transition counts. In other words, the string of 1's and 0's which represent the observed purchasing behavior of a particular family, is converted into a series of transition counts, f_{ij} , where $i, j = 0, 1$, and i is the state of the system prior to transition and j is the state of the system after transition. For instance, the purchase sequence, 1001101011, is represented as: $f_{11} = 2$ (indicating two "transitions" from state 1 to state 1), $f_{10} = 3$ (indicating three transitions from state 1 to state 0), $f_{01} = 3$, and $f_{00} = 1$.

We begin by defining the matrix beta distribution. Let $\tilde{P} = [\tilde{p}_{ij}]$ be an $(N \times N)$ stochastic matrix and let $M = [m_{ij}]$ be an $(N \times N)$ matrix of parameters. Then if \tilde{P} has the matrix beta distribution with parameter M , the joint density is

$$f_{MB}(P|M) = \prod_{i=1}^N \beta_N(M_i) \prod_{j=1}^N p_{ij}^{m_{ij}-1}$$

where

$$\beta_N(M_i) = \frac{\Gamma(M_i)}{\prod_{j=1}^N \Gamma(m_{ij})}, \quad M_i = \sum_{j=1}^N m_{ij}, \quad \Gamma \text{ is the gamma function,}$$

and $\sum_{j=1}^N p_{ij} = 1$.

The mean of \tilde{p}_{ij} is given by

$$E(\tilde{p}_{ij}) = \frac{m_{ij}}{M_i}, \quad i, j = 1, \dots, N .$$

The variance of \tilde{p}_{ij} is

$$\text{Var}(\tilde{p}_{ij}) = \frac{m_{ij}(M_i - m_{ij})}{M_i^2(M_i + 1)}, \quad i, j = 1, \dots, N$$

and the covariance of $\tilde{p}_{ij}, \tilde{p}_{k\ell}$ is

$$\begin{aligned} \text{Cov}(\tilde{p}_{ij}, \tilde{p}_{k\ell}) &= \frac{-m_{ij}m_{i\ell}}{(M_i)^2(M_i+1)} && \text{for } i = k = 1, \dots, N \\ & && j, \ell = 1, \dots, N \\ &= 0 && i \neq j \end{aligned}$$

\tilde{P} represents the transition matrix, P , of the two-state Markov chain and M indicates our prior judgments about \tilde{P} . A method of determining M is discussed in the body of the paper.

The likelihood function for the transition counts is given by the Whittle distribution. Let \tilde{F} be an $(N \times N)$ stochastic matrix with elements \tilde{f}_{ij} . $P = [p_{ij}]$ is the $(N \times N)$ matrix of transition probabilities. u is the state of the system prior to transition and v is the final state. Then, the Whittle distribution is

$$f_W(F|u, v, P) = F_{vu}^* \frac{\prod_{i=1}^N \Gamma(f_{i.} + 1)}{\prod_{i=1}^N \prod_{j=1}^N \Gamma(f_{ij} + 1)} \prod_{i=1}^N \prod_{j=1}^N p_{ij}^{f_{ij}}$$

where $f_{i.} = \sum_{j=1}^N f_{ij}$ and F_{vu}^* is defined as follows. $F^* = [f_{ij}^*]$ is an $(N \times N)$ matrix defined by

$$\begin{aligned} f_{ij}^* &= \delta_{ij} - \frac{f_{ij}}{f_{i.}} && f_{i.} > 0 \\ &= \delta_{ij} && f_{i.} = 0 \end{aligned}$$

F_{vu}^* is the (v, u) th cofactor of F^* and $\delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$

The predictive distribution for \tilde{F} , the matrix of transition counts, is a beta-Whittle distribution with parameter matrix, M . It is important to note that the predictive distribution for F does not depend upon P . We derive it as follows:

$$\begin{aligned} f_{\beta W}(F|u, v, M) &= \int_P f_W(F|u, v, P) f_{MB}(P|M) dP \\ &= \int_P \left[\frac{F_{vu}^* \prod_{i=1}^N \Gamma(f_{i.} + 1)}{\prod_{i=1}^N \prod_{j=1}^N \Gamma(f_{ij} + 1)} \prod_{i=1}^N \prod_{j=1}^N p_{ij}^{f_{ij}} \right] \left(\prod_{i=1}^N \beta_N(M_i) \prod_{j=1}^N p_{ij}^{m_{ij}-1} \right) dP \end{aligned}$$

$$= F_{vu}^* \frac{\prod_{i=1}^N \Gamma(f_{i\cdot} + 1) \beta_N(M_i)}{\prod_{i=1}^N \Gamma(f_{ij} + 1)} \int_P \prod_{j=1}^N p_{ij}^{f_{ij} + m_{ij} - 1} dp$$

$$f_{\beta W} (F|u, v, M) = F_{vu}^* \frac{\prod_{i=1}^N \Gamma(f_{i\cdot} + 1)}{\prod_{i=1}^N \Gamma(f_{ij} + 1)} \frac{\prod_{i=1}^N \beta_N(M_i)}{\prod_{i=1}^N \beta_N(S_i)}$$

where $s_{ij} = m_{ij} + f_{ij}$ and $S_i = \sum_{j=1}^N s_{ij}$.

The Predictive Distribution for a Bernoulli Process:
The Beta-Binomial Distribution

For the Bernoulli process, we use the beta distribution as our prior for \tilde{P} , the probability of a success, and the binomial distribution as the likelihood function for r , the number of successes we observe in our sample. The beta distribution is defined by

$$f_{\beta}(p|r', n') = \frac{1}{\beta(r', n'-r')} p^{r'-1} (1-p)^{n'-r'-1} \quad 0 < p < 1$$

where $\beta(r', n'-r') = \frac{\Gamma(n')}{\Gamma(r')\Gamma(n'-r')}$ and r' and n' are parameters with $n', r' > -1$.

The likelihood function is a binomial distribution

$$f_b(r|n, p) = \binom{n}{r} p^r (1-p)^{n-r} \quad r = 0, 1, \dots, n$$

where $\binom{n}{r} = \frac{n!}{r!(n-r)!}$.

Then our predictive for r becomes a beta-binomial and is derived as follows:

$$\begin{aligned} f_{\beta b}(r|n, r', n') &= \int_0^1 \binom{n}{r} p^r (1-p)^{n-r} \cdot \frac{1}{\beta(r', n'-r')} \cdot p^{r'-1} (1-p)^{n'-r'-1} dp \\ &= \frac{\binom{n}{r} \beta(r+r', n+n'-r-r')}{\beta(r', n'-r')} \int_0^1 \frac{p^{r+r'-1} (1-p)^{n+n'-r-r'-1}}{\beta(r+r', n+n'-r-r')} dp \\ &= \binom{n}{r} \frac{\beta(r+r', n+n'-r-r')}{\beta(r', n'-r')} \end{aligned}$$

References

- Aaker, D. A. A new method for evaluating stochastic models of brand choice. Journal of Marketing Research, 1970, 7, 300-306.
- Aaker, D. A. The New Trier stochastic model of brand choice. Management Science, 1971, 17, B435-450.
- DeGroot, M. H. Optimal statistical decisions. Hightstown, N.J.: McGraw-Hill, 1970.
- Jeffreys, H. Theory of probability. (2nd ed.) London: Oxford Press, 1948, Chapter V.
- Jones, J. M. A comparison of three models of brand choice. Journal of Marketing Research, 1970, 7, 466-473.
- Jones, J. M. A stochastic model for adaptive behavior in a dynamic situation. Management Science, 1971, 17, 484-497.
- Martin, J. J. Bayesian decision problems and Markov chains. New York: Wiley, 1967.
- Massy, W. F., Montgomery, D. B., & Morrison, D. G. Stochastic models of buying behavior. Cambridge, Mass.: M.I.T. Press, 1970.
- Morrison, D. G. Testing brand-switching models. Journal of Marketing Research, 1966, 3, 401-409.
- Rao, C. R. A study of large sample test criteria through properties of efficient estimates. Sankhyā, 1961, A23, 25-40.

DEMOGRAPHIC SEGMENTATION OF LONG DISTANCE BEHAVIOR:

DATA ANALYSIS & INDUCTIVE MODEL BUILDING¹

A. Marvin Roscoe, Jr.²
 American Telephone & Telegraph Company

and

Jagdish N. Sheth
 University of Illinois

The objective of the study was to explore the appropriateness of several statistical techniques in developing predictive models of consumers' long distance telephone expenditure based on the analysis of socioeconomic and demographic characteristics. Specifically, the paper examines the relative efficacy of stepwise multiple regression, monotonic AID (Automatic Interaction Detector) and free AID in analyzing large scale data.

Description of MRIS Data Bank

Most consumer behavior research to date has been ad hoc, fragmentary, and exploratory. Only recently have large corporations begun to generate continuous and systematic information about the market place as part of their marketing information systems. The Bell System Companies provide communication services in the 48 continental states and the District of Columbia to 104 million telephones, 83% of the total telephones in the United States. The need to understand this enormous market is self-evident, not only from a traditional marketing view but also from the social and economic considerations inherent in the management of a regulated utility. To help meet this need, a large scale Market Research Information System (MRIS) has been established, consisting of a national longitudinal panel of some 60,000 customers representing both the business and residence markets.

MRIS panel members were selected by multistage stratified sampling procedures from the customer files of each of the one hundred accounting offices of the Bell System where customer billing is performed. A panel of 600 customers, evenly divided between business and residence customers, was selected from each of these accounting offices. Currently, the MRIS data bank contains more than 126 million card image records, and is growing at the rate of 3 1/2 million records each month. The MRIS panel excludes certain types of accounts, such as the U.S. Government, which are handled separately from a communications view point, and certain specialized types of communications services such as private line and data services.

For each panel member, the MRIS data bank stores the following information:

1. A basic equipment record consisting of service and equipment data, such as the number and type of telephone lines, number of extension phones and other vertical (optional) services including Princess* and Trimline* phones, Touch-Tone* service and additional Directory listings. These data are updated whenever panel members change their service or equipment.

*/Registered trademark of the Bell System.

2. A billing amount record listing charges for local service, additional message units (where applicable), a summary of long distance billing, taxes and other charges or credits as shown on the customer's bill. This record is expanded every month.
3. A long distance record listing billing details for each message found on the customer's billing statement, such as the date and time of the call, type of call (direct dial or operator handled), length of conversation and amount of charge. This record is also expanded each month.
4. A demographic record containing a socioeconomic and demographic profile of the residence customer's household unit. These data have been obtained from a mail questionnaire, and include age, sex, education and occupation of the head of household, family size and composition, its mobility characteristics and family income. The residence profile is updated every three years, and consideration is presently being given to collecting additional information on the residence customer's fundamental value system as well as his generalized attitudes toward the telephone as a means of communication.

To be able to comprehend this enormous amount of information at the microlevel of the individual customer, basic research is underway to develop analytic strategies in the interest of building predictive micro and macro models of buyer behavior for the telephone industry.

This paper represents one of our projects designed to develop an understanding of the long distance telephone behavior of the residence customer. The relationship of socioeconomic and demographic factors to long distance behavior is especially important to insure that the rate structures filed by the Telephone Companies and approved by the regulatory agencies are equitable to the various socioeconomic customer segments in the country³.

In order to examine the characteristics of the data and the associated problems in their analysis, the study was limited to the 793 panel customers from two Northeastern states. The two groups of customers were chosen on the basis of comparability of long distance expenditure, providing a sample size that would not unduly favor one particular analytic technique. The focus of this paper is, however, data analysis and not inference, although decision and predictive models are being built based on the findings from this type of data analysis utilizing larger and more generalized samples of the population.

Table 1 lists the fourteen socioeconomic and demographic variables and the dependent variable, long distance expenditure⁴, with their means and standard deviations. To avoid the effects of seasonality and holidays, the long distance behavior variable was based on the monthly average of a year's history for each residence customer, expressed in dollars and cents. The dollar signs have been omitted for the sake of simplicity. Among the fourteen socioeconomic and demographic variables, two index variables have been included, the socioeconomic status (SES) index and the Life Cycle index. The SES index is a score developed from a composite of the education and occupation of the head of household and family income level using the procedures of the U.S. Bureau of the Census (1963). The Life Cycle index is determined from the age and marital status of the head of household and family composition; following the procedures used by the Survey Research Center at the University of Michigan (Lansing & Kish, 1957).

TABLE 1
List of Variables

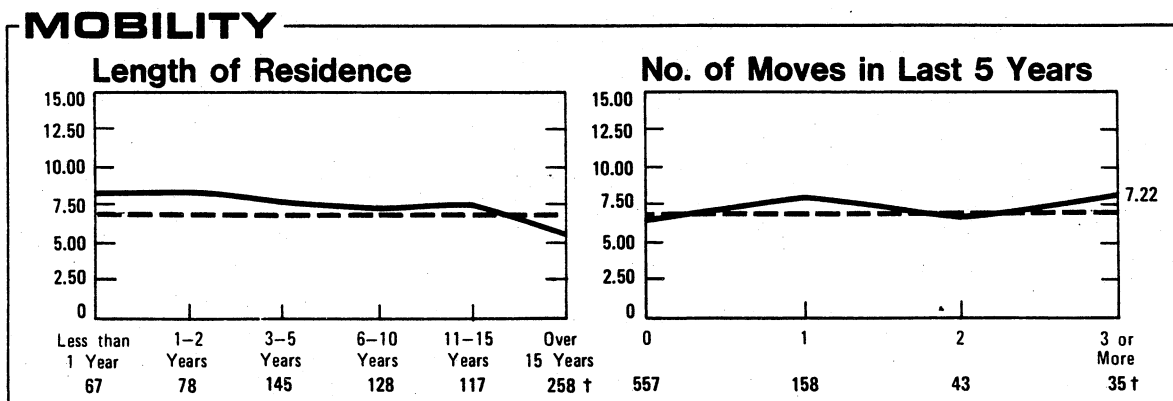
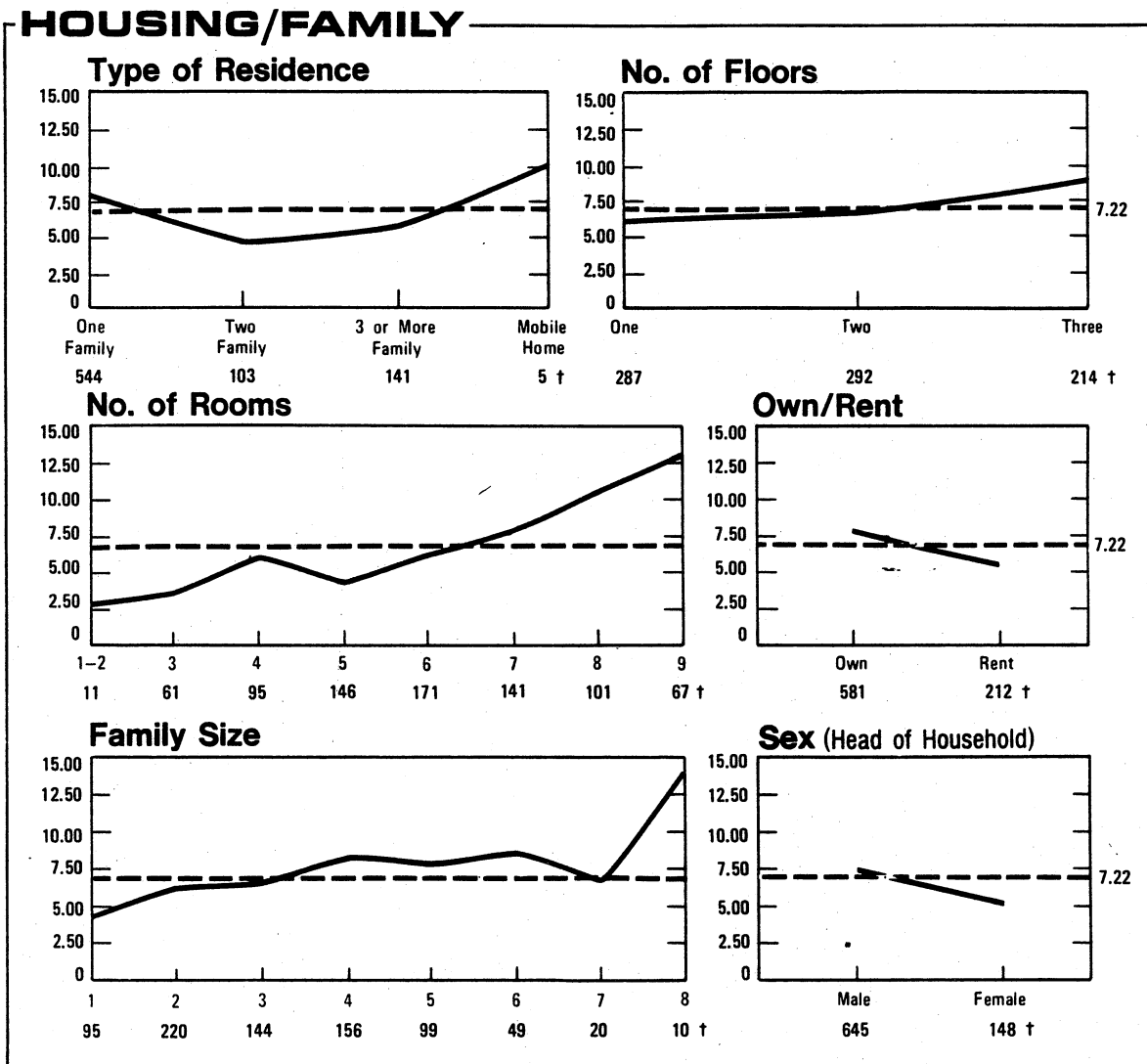
Number	Description	Mean	Standard Deviation
1	Socio-Economic Status	6.385	2.079
2	Own/Rent	1.267	0.443
3	Type of Residence	1.504	0.803
4	No. of Floors	1.908	0.790
5	No. of Rooms	5.974	1.747
6	Length of Residence	4.165	1.662
7	No. of Moves (in past 5 yrs.)	1.440	0.788
8	Sex of H. H.	1.187	0.390
9	Age of H. H.	5.009	1.383
10	Occupation of H. H.	6.095	2.322
11	Education of H. H.	4.276	1.789
12	Family Income	4.511	1.801
13	Family Size	3.279	1.629
14	Life Cycle	4.764	1.585
15	Long Distance Expenditure (average month)	7.219	10.376

Problems of Data Analysis and Alternative Statistical Approaches

In Figure 1, long distance expenditure is plotted for each socioeconomic and demographic variable. The variables have been grouped in four categories derived from a prior factor analysis of these data. Examining the plots clearly points out the following data problems typical of most survey research (Morgan & Sonquist, 1963; Carman, 1967; Sonquist, 1970):

1. All of the demographic variables are discrete rather than continuous, although many of them do have successive class intervals containing large numbers of observations.
2. The variables have a mixture of scales consisting of nominal and interval-scaled data.
3. The relationship of long distance expenditure with many of the demographic variables is not linear.
4. In some cases, the relationship is not even monotonic.
5. The demographic variables may be related to long distance telephone behavior in an interactive manner rather than in a simple additive manner.
6. The demographic variables tend to be correlated with one another, which may be a serious problem when using regression analysis (Blalock, 1963). Table 2 summarizes some of the highly multicollinear variables.

Figure 1
Average Long Distance Expenditure
By Demographic Category



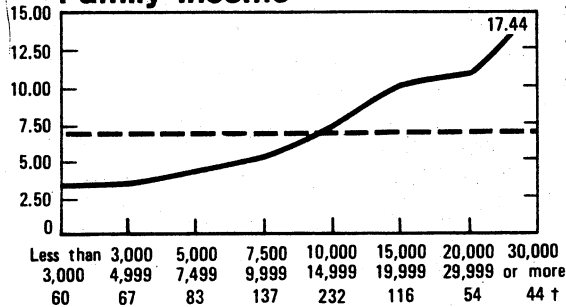
† Number of Cases in Each Category

All Customer Average - - - -

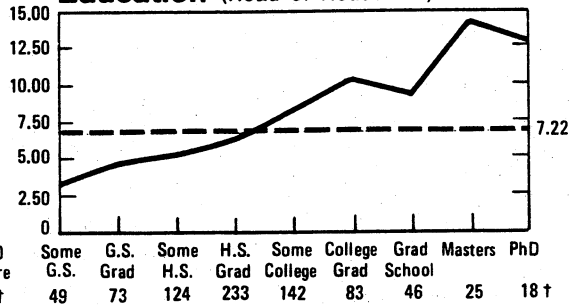
Figure 1 (cont'd)
Average Long Distance Expenditure
By Demographic Category

SOCIO-ECONOMIC

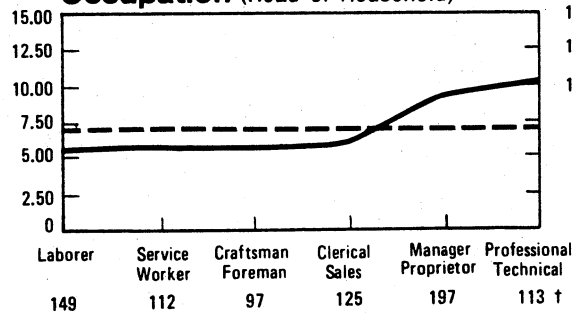
Family Income



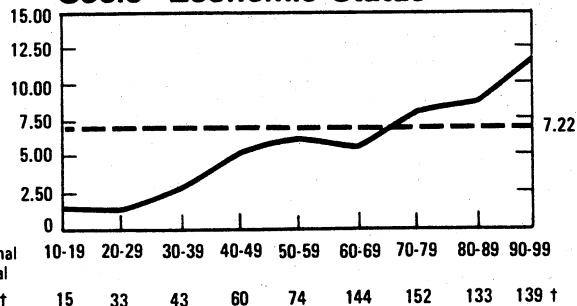
Education (Head of Household)



Occupation (Head of Household)

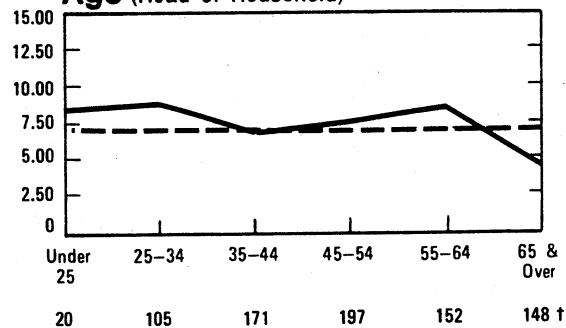


Socio - Economic Status*

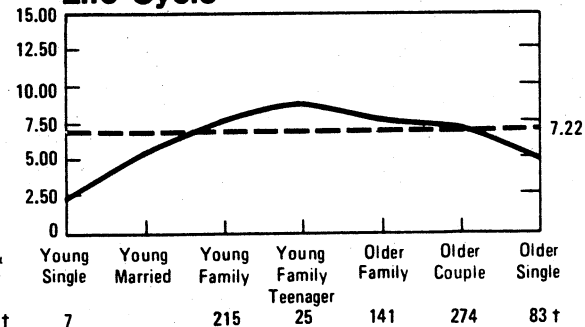


LIFE CYCLE

Age (Head of Household)



Life Cycle



* Based upon U.S. Bureau of Census, Methodology and Scores of Socioeconomic Status, Working Paper No. 15, 1963

All Customer Average -----

† Number of Cases in Each Category

TABLE 2

Multicollinearity Among Demographic Variables

Variable	Correlation	Variable
1. SES Score	0.75 0.76 0.78	Occupation of H. H. Education of H. H. Income of H. H.
2. Own/Rent	0.69 -0.56	Type of Residence No. of Rooms
3. Type of Residence	0.54	No. of Rooms
4. No. of Floors	0.52	No. of Rooms
5. Length of Residence	-0.71 0.50	No. of Moves Age of H. H.
6. Age of H. H.	0.81	Life Cycle
7. Occupation of H. H.	0.51	Family Income
8. Education of H. H.	0.51	Family Income

Starting with a large body of empirical data and little or no prior theory with which to develop functional relationships, it is difficult to construct a multivariate model without first investigating the effects of these data problems as they relate to analysis by standard statistical methods. In fact, the authors believe that the blind use of a statistical method may produce great harm from misinterpretation of the data, and could even result in throwing out important data as irrelevant or useless for a marketing problem.

Our objective was, therefore, to first explore several analytic strategies which take into consideration, to varying degrees, these data problems. Stepwise multiple regression, AID with the monotonic restriction on the predictor variables, and AID without the monotonic restriction were chosen as the three alternative strategies because each responds somewhat differently to these data characteristics. Such a combination should, therefore, both avoid the problem of forming an unsupportable advance hypothesis and give the analyst perspective on the actual structure under observation.

Multiple regression is a robust method, rich in both data analysis and inference. In addition, regression analysis, with some variations, can take into account the problem of mixed scales and class interval data. For example, by using dummy variables it is possible to include a number of nominally scaled demographic descriptors such as ownership of residence and sex of the head of household. Finally, stepwise multiple regression considers the problem of multicollinearity by developing partial correlations at each step,

thereby eliminating variables that are highly correlated with variables already included in the regression equation.⁵ However, since the regression equation is a linear additive model, it is not capable of effectively handling the problems of nonlinear, nonmonotonic and interactive relationships⁶.

The objective in AID analysis is to partition the total sample into an optimal set of nonoverlapping subgroups, developed from the profiles of the predictor variables, whose categories explain more of the variation in the dependent variable than do any other set of subgroups. This objective is achieved by a sequential partitioning of the total sample into two subgroups based on the split of a single predictor variable which produces the largest ratio of between sum of squares to total sum of squares. This process is repeated on each of the subgroups until some minimum level of explained variance is encountered or a minimum sample size is reached in the subgroup. Thus, one-way analysis of variance is explicitly included in the analysis.

Because the splitting of groups is sequential, the AID analysis is a stepwise procedure similar to the stepwise regression method, and so minimizes the problem of multicollinearity. However, the optimal split at each step is not based on a predictor's contribution to reducing the error variance in the total sample but the variance in the subgroup.

Several researchers have used the AID technique in marketing, either for developing segments (Assael, 1970) or for model building (Carman, 1967; Armstrong & Andress, 1970). The technique itself is described by Morgan and Sonquist (1963), Sonquist and Morgan (1964), and Sonquist (1970). The AID program is capable of handling both the categorical and class interval predictor variables, regardless of whether their relationship is linear, nonlinear or nonmonotonic with respect to the criterion variable. Of course, the criterion, or dependent, variable may be continuous and in the case of long distance expenditure it is. Finally, and most importantly, this procedure is capable of handling both the additive and the interactive relationships of a set of predictors with the criterion variable.

Two types of AID analyses were used to separately examine the nonmonotonic and the interactive effects of the relationships between long distance expenditure and the demographic variables. The first, monotonic AID analysis, preserves the ordinality of the predictor variable when it is chosen as a candidate to split the sample. Therefore, the two new subgroups are defined as above and below the boundary of a category interval of the predictor variable. For example, given the eight categories of income, there are only seven (K-1) comparisons possible by splitting the group at each of the adjacent categories. By definition, therefore, monotonic AID analysis is capable of handling nonlinear relationships as long as they are monotonic or order-preserving.

The second procedure, free AID analysis, allows the split on a predictor variable without regard to the order of the categories of that variable. Thus, there is a much larger number of combinations of the predictor variable categories which may be examined to split the sample. This removes the nonlinear restriction and allows for the analysis of a nonmonotonic relationship if one exists between the predictor and the criterion variable.

Comparative Data Analysis and Results

The stepwise linear regression analysis was performed using the UCLA Biomedical computer program BMD 02R (Dixon, 1971). To avoid highly collinear variables and random effects, an F value of 3.85, comparable to 0.05 level of significance, was set for a predictor variable to enter into the equation. The results of the stepwise regression analysis are summarized in Table 3. The multiple R was 0.36, resulting in 12.68 percent of the variance in long distance expenditure being explained by four of the demographic predictors. These four significant predictors and their associated explained variances are (1) family income, 9.86% (2) number of rooms, 0.75% (3) length of residence, 0.98% and (4) life cycle of the family 1.09%. The relationship is positive with income, number of rooms and life cycle, but negative with length of residence. In short, the greater the income, the more rooms in the residence unit, the later the stage of the life cycle, and the more recent the move of a residence customer, the greater the average long distance expenditure will be. It is interesting to note that life cycle as an index variable performed better than its component demographic variables but the SES index did not perform better than income.

TABLE 3

Stepwise Linear Regression Analysis

Variables in Equation						
Step	Variable	Beta Coef.	Standard Error	*Multiple R	RSQ	F Ratio to Enter
1	(Constant	0.0001)				
	Income	0.2699	0.0387	0.3140	0.0986	86.50
2	No. of Rooms	0.1385	0.0393	0.3257	0.1061	6.62
3	Length of Residence	-0.1613	0.0391	0.3404	0.1159	8.74
4	Life Cycle	0.1224	0.0389	0.3561	0.1268	9.88

Analysis of Variance

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	F Ratio
Total	792	85,264	2,703.05	28.61*
Regression	4	10,812	94.48	
Residual	788	74,452		

Percent Variance Explained 12.68

*Significant at the 0.01 level.

TABLE 4
Monotonic AID Analysis

		Group Definitions					Splitting Definitions				
Group Split	Group Number	Size	Mean Value	Standard Deviation	Total Sum of Squares	Predictor Variable	Variable Values	Between Sum of Squares	Percent Variance Explained	t Value	
	1	793	7.22	10.37	85,264	Family Income	≤ \$15,000	6,537	7.67	8.10	
1	2	579	5.47	7.06	28,854	Family Income	> \$15,000				
3	3	214	11.94	15.26	49,873	Number of Rooms	1 to 7	1,918	2.25	2.91	
3	4	120	9.29	9.87	11,696	Number of Rooms	8 & 9				
5	5	94	15.33	19.64	36,259	Length of Residence	1-10 yrs.	960	1.12	1.58	
5	6*	49	18.39	25.14	30,976	Length of Residence	11+ yrs.				
7*	7*	45	11.99	9.80	4,323	Family Income	< \$10,000	1,390	1.63	5.40	
8	8	347	4.21	5.76	11,498	Family Income	\$10,000-\$15,000				
9	9	232	7.37	8.29	15,966	Number of Rooms	1 to 5	377	0.44	2.36	
9	10	82	5.64	6.56	3,532	Number of Rooms	6 to 9				
11	11	150	8.31	8.96	12,057	Number of Moves	0 & 1	141	0.17	1.32	
11	12	114	7.77	8.29	7,840	Number of Moves	2 or more				
13*	13*	36	10.04	10.64	4,076	Education of H. H.	To College Deg. Graduate School	392	0.46	2.02	
4	14	90	8.25	8.66	6,747	SES Score	10 - 39	464	0.54	3.81	
8	15*	30	12.42	12.32	4,557	SES Score	40 - 99				
17	16*	91	2.27	3.88	1,368	Length of Residence	1-5 yrs.	255	0.30	2.63	
17	17	256	4.90	6.14	9,666	Length of Residence	6+ yrs.				
17	18	104	6.10	7.49	5,831	Life Cycle of Family	1 to 5	130	0.15	1.37	
19*	19*	152	4.07	4.35	3,580	Life Cycle of Family	6, 7				
12	20*	64	6.82	7.54	3,641	Age of H. H.	18 - 54	227	0.27	1.75	
14	21*	50	8.97	9.02	4,069	Age of H. H.	55+				
14	22*	55	6.98	6.49	2,314	SES Score	40 - 69	240	0.28	2.09	
23*	23*	35	10.24	10.96	4,206	SES Score	70 - 99				
18	24	71	7.14	8.47	5,102	Life Cycle of Family	1 - 3	201	0.24	1.68	
18	25*	33	3.88	3.85	489	Life Cycle of Family	4 - 7				
24	26*	36	5.48	6.39	1,472	SES Score	10 - 79	442	0.52	3.39	
27*	27*	35	8.85	9.90	3,429	SES Score	80 - 99				
10	28*	51	3.83	3.74	714						
29*	29*	31	8.62	8.75	2,376						
								13,674	16.04		

*Final Groups

Several questions are implicit in these findings. First, the bulk of the variance explained is concentrated in income (77.7%) with relatively little contribution from the other demographic variables. Secondly, the total amount of explained variance is relatively lower than might be expected⁷. And third, most other demographic variables fail to exhibit any relationship on a partial correlation basis. There are obviously two answers. First, no relationship may exist with these other variables, so that long distance expenditure is determined by other factors not in the equation. However, and secondly, it is possible that the linear additive model built into the regression suppresses any non-linear or interactive relationship between the demographic variables and long distance telephone behavior. Unless the latter explanation is ruled out, good data may be discarded due to inappropriate analytic methods.

The monotonic AID analysis was the second method used. This procedure allows for monotonic, nonlinear and interactive relationships between the predictor variables and the criterion variable as noted. To avoid unstable results and to meet sampling error requirements, the AID analysis was based on the additional constraints of a minimum sample size of 30 in each final subgroup, and a minimum percent variance explained equal to or greater than 0.6 percent at each step.

The statistical results are summarized in Tables 4 and 5. The explained variance was increased from 12.68 percent, using regression analysis, to 16.04 percent using AID and allowing monotonic and interactive relationships. Table 4 shows that the SES score, the education and age of the head of household and number of moves have entered into the analysis. The additional explanatory power comes from (1) the inclusion of these variables and (2) the increases in the predictive power of the variables as against the regression equation. The best examples of increased predictive power are the SES score, which was 1.34%, and the number of rooms which increased from 0.75% to 2.69% variance explained. These values are the summation of individual percent variance explained in Table 4. This increased predictive power can be explained by the fact that both of these variables have a step function with the long distance expenditure as seen from the plots in Figure 1. A similar step function in length of residence also slightly increases its predictive power.

TABLE 5

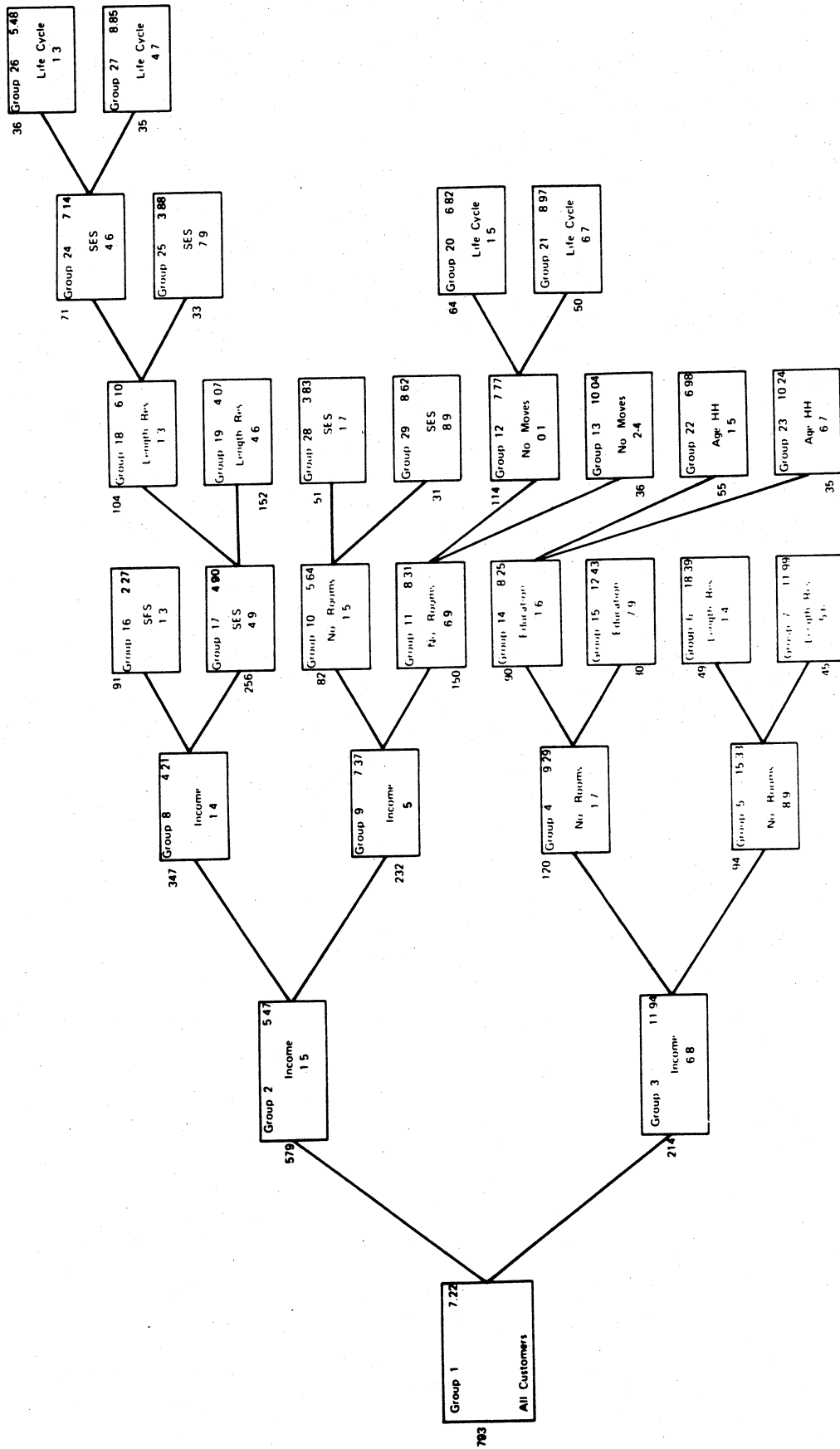
Monotonic AID Analysis

Analysis of Variance				
Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	F Ratio
Total	792	85,264		
Between	15	13,674	911.61	9.89*
Within	777	71,590	92.14	

Percent variance explained 16.04

*Significant at the 0.01 level

FIGURE 2
Monotonic AID Analysis



On the other hand, the predictive power of income and life cycle decreased slightly in monotonic AID. This is largely due to the small number of cases at the upper end of the income scale and at the lower end of the life cycle index. These small cell sizes do not permit further subdivision due to the restriction of the minimum size in the final groups formed.

A careful examination of the group splits which result in large between sum of squares reveals that they are directly a function of the truncation in the monotonic relationship of the predictor variable with the criterion variable. Thus, the greater the rise of the step in the function between the predictor and the criterion variable, the greater the relative predictive power that variable possesses. Surprisingly, the interaction among the significant demographic variables is not as great as expected. This is shown in the tree diagram of Figure 2 by the relatively good symmetry of splits where the predictive variables appear on both branches of the split. Stronger interaction between the demographic and the socioeconomic variables had been anticipated; however, this interaction does not seem to be present in the data.

Finally, a very important benefit of monotonic AID analysis comes from the fact that it matches the managerial problem definition and decision-making process. Typically, marketing management is interested in market differentiation and discrimination for better marketing effectiveness. Furthermore, the market differentiation is based on segmenting customers who are presumed to have different wants and desires. The demographic variables are considered the most common casual factors in bringing these differences to light, especially by the regulatory and other governmental agencies in the utility industry. The AID results have been represented as a diagram in Figure 2 which is both meaningful and communicable to management. For example, it indicates that customers with income above \$15,000, with residence units consisting of eight or more rooms, and with less than ten years of residence at their present location have the highest average of long distance calling and expenditure (group six). On the other hand, customers with less than \$10,000 income and an extremely low SES score manifest the lowest amount of average monthly expenditure (group sixteen). Both of these extreme groups, as well as the other segments, are meaningful and relate to management's prior experiences and decisions. In view of the fact that half the problem in successful marketing research is its effective communication, AID seems to be an advantageous analytical strategy.

The third analytic technique is free AID analysis where the nonmonotonicity of the predictor-criterion relationship is taken into account in addition to the nonlinearity and interactive aspects. The same stopping criteria were utilized here (minimum subgroup size greater than or equal to 30 and 0.6 percent variance explained at each step). The statistical results are summarized in Tables 6 and 7 and Figure 3.

Free AID analysis increases the predictive power of the demographic variables from 16.04 percent to 20.23 percent when compared to monotonic AID analysis. In the process, it includes occupation and type of residence variables; however, the bulk of the increased predictive power in this analysis comes from the demographic variable age of head of household, which has the greatest nonmonotonic relationship with long distance expenditure. Somewhat smaller increases in the predictive power of number of rooms, education, and life cycle are also due to their nonmonotonic relationship with long distance expenditure.

TABLE 6
Free AID Analysis

Group Definitions			Splitting Definitions							
Group Split	Group Number	Size	Mean Value	Standard Deviation	Total Sum of Squares	Predictor Variable	Variable Values	Between Sum of Squares	Percent Variance Explained	t Value
1	1	793	7.22	10.37	85,264	Family Income	1 to 5 6 to 8	6,537	7.67	8.10
2	2	579	5.47	7.06	28,854	Number of Rooms	2,4 to 7 3,8,9	2,024	2.37	2.99
3	3	214	11.94	15.27	49,873	Age of H. H.	4,5,7 3,6	2,980	3.49	2.90
4	4	117	9.14	19.42	36,591	Family Income	1 to 4 5	1,390	1.63	5.40
5	5	97	15.32	28.91	25,918	Occupation of H. H.	6 3,4,7-9	388	0.46	2.39
6	6	66	11.52	5.76	11,498	Type of Residence	2,4 1,3	443	0.52	2.53
7*	7*	31	23.41	8.30	15,966	Education of H. H.	1,3-5,7-8 2,6,9	214	0.25	1.66
8	8	347	4.21	7.33	12,893	SES Score	1-3,9	488	0.57	3.91
9	9	232	7.37	8.42	2,685	Education of H. H.	1-3,5,6 4,7-9	526	0.62	2.37
10*	10*	50	4.90	7.33	12,893	Length of Residence	2,4-6 1,3	414	0.49	3.36
11	11	182	8.05	4.41	585	Life Cycle of Family	2 to 4 5 to 7	445	0.52	1.98
12*	12*	30	4.53	8.84	11,865	Life Cycle of Family	1 to 5 6,7	168	0.20	1.66
13	13	152	8.74	7.86	7,529	Age of H. H.	3,4 2,5-7	447	0.53	2.54
14	14	122	8.15	11.72	4,122	Number of Rooms	2,3,5,6 4,7-9	360	0.42	2.59
15*	15*	30	11.13	3.84	1,370	Occupation of H. H.	3,4,6,7-9 8	216	0.25	3.20
16	16	93	2.25	6.16	9,640	Number of Rooms	4-6 7	100	0.12	1.41
17	17	254	4.92	7.09	3,422	Length of Residence	1,2,4,6 3,5	106	0.12	2.76
18	18	68	7.34	12.21	7,310	SES Score				
19*	19*	49	11.64	4.70	4,060	Education of H. H.				
20	20	184	4.14	8.59	5,166	Length of Residence				
21	21	70	6.99	4.70	3,213	Life Cycle of Family				
22*	22*	31	8.77	10.18	4,035	Life Cycle of Family				
23*	23*	35	13.97	10.74	4,425	Age of H. H.				
24	24	78	7.27	7.53	2,936	Number of Rooms				
25*	25*	44	9.71	8.17	663	Occupation of H. H.				
26*	26*	37	4.61	4.23	4,056	Number of Rooms				
27*	27*	33	9.67	11.09	442	Length of Residence				
28*	28*	32	4.69	3.72	3,623	Life Cycle of Family				
29*	29*	46	9.06	8.88	2,110	Age of H. H.				
30*	30*	152	3.64	3.73	1,734	Number of Rooms				
31*	31*	32	6.50	7.36	1,020	Occupation of H. H.				
32*	32*	36	6.20	5.32	2,302	Length of Residence				
33*	33*	32	8.63	8.48	285					
34*	34*	61	1.47	2.16	979					
35*	35*	32	3.72	5.53						
								17,246	20.23	

* Final Groups

TABLE 7

Free AID Analysis

Analysis of Variance				
Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	F Ratio
Total	792	85,264		
Between	18	17,246	958.07	10.90*
Within	774	68,018	87.88	

Percent variance explained 20.23

*Significant at the 0.01 level

Surprisingly, the predictive power of both the SES index and length of residence decreased in the free AID analysis. This is attributable to the extreme skewness of the two predictor variables.

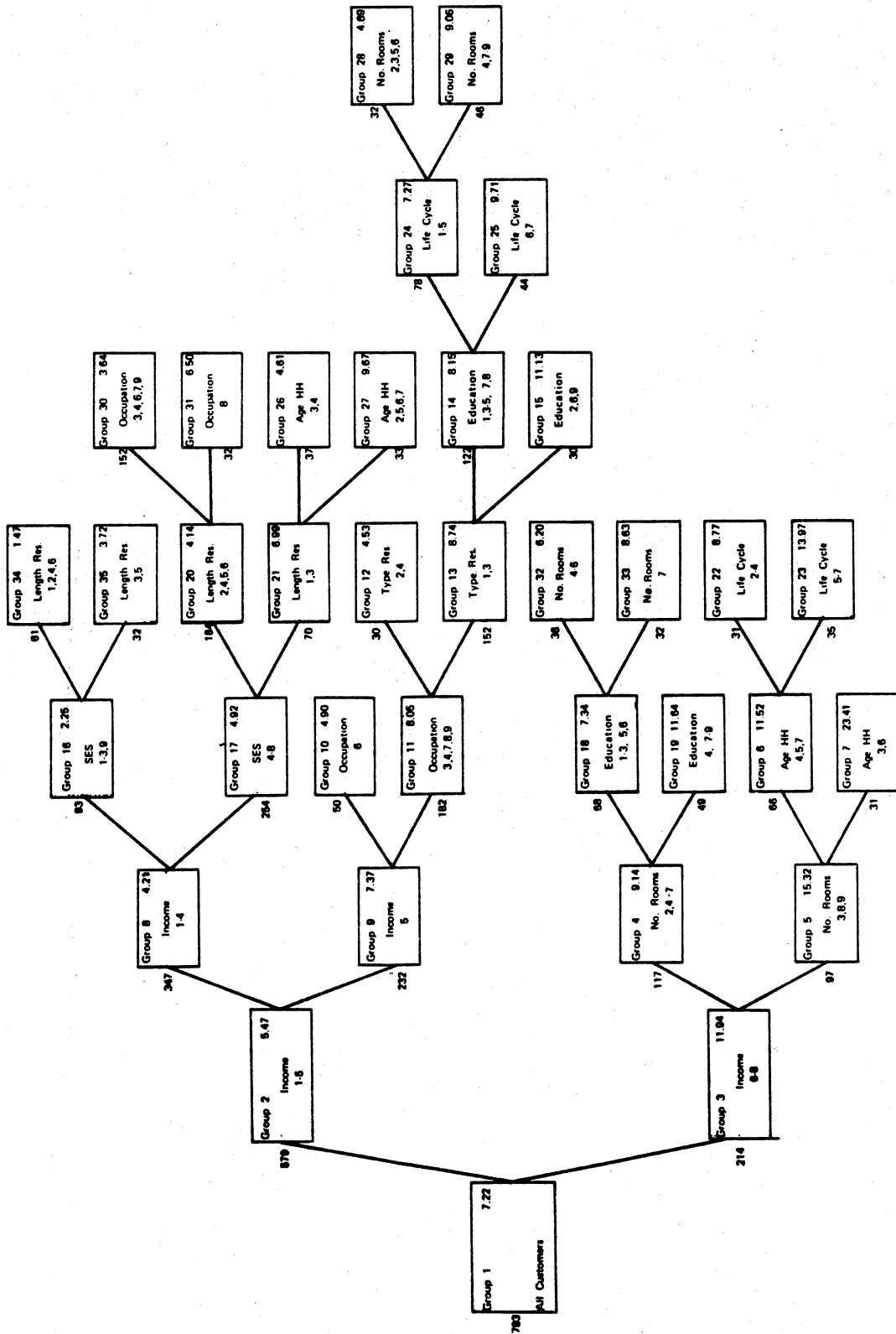
The free AID analysis reveals some subtle differences among customer segments which are hidden in the monotonic AID analysis. For example, the seventh group which has the highest average monthly bill, consists of customers who have greater than \$15,000 income, both small and large residence units (three rooms and eight or more rooms) and who are both relatively young and relatively old (between 25 and 34 years and between 55 and 64 years). This was not fully revealed either in monotonic AID or in stepwise regression.

TABLE 8

Comparative Analysis of Predictor Variables

Percent Variance Explained			
	Stepwise Regression	Monotonic AID	Free AID
Family Income	9.86	9.30	9.30
No. of Rooms	0.75	2.69	2.91
Length of Residence	0.98	1.42	0.61
Life Cycle	1.09	0.39	0.72
SES Score	--	1.34	0.57
Age of H. H.	--	0.27	4.02
No. of Moves	--	0.17	--
Education of H. H.	--	0.46	0.87
Occupation of H. H.	--	--	0.71
Type of Residence	--	--	0.52
	<u>12.68</u>	<u>16.64</u>	<u>20.23</u>

FIGURE 3
Free AID Analysis



In view of the fact that the demographic variables which explain the variance in long distance expenditure differ for the three analytic methods, Table 8 has been prepared to summarize their relative contribution toward explaining variance in the criterion variable. Income, which was by far the best predictor, does not do as well in the AID analyses, primarily due to the problem of sample size in the extreme cells. The other three demographic variables - number of rooms, SES score and length of residence - did very well in monotonic AID due to two factors: (1) they have a step function relationship with the criterion variable, and (2) the distributions are badly skewed. Finally, age of head of household does best in free AID primarily due to its nonmonotonic relationship with the criterion variable.

Having removed the nonlinear and nonmonotonic constraints and permitted interactive relationships in the analysis, the explained variance has increased from 12.68 to 20.32 percent. However, the unexplained variance still remains quite large. An additional advantage of the AID analysis is the ability to investigate the variance structure by customer segment. Monotonic AID and the Free AID analyses both developed three branches defined by low, medium and high income categories. These represent 44%, 29% and 27% of the sample size. Table 9 summarizes the effects for each branch.

TABLE 9

Summary of Variance
(in Percentages)

Income Category	Variance in each branch	Monotonic AID		Free AID	
		Explained	Unexplained	Explained	Unexplained
Less than \$10,000	13.48	1.36	12.12	1.96	11.52
\$10,000 to \$15,000	18.73	1.28	17.45	1.85	16.88
More than \$15,000	58.49	4.10	54.39	7.12	51.37

Note: 9.30% is explained by the three income categories.

Finally, noting that the bulk of the unexplained variance is in the high income branch a search of Table 4 and 6 shows that final group 6 in the monotonic AID analysis has 36+ percent of the total variance, and group 7 in free AID analysis has 30+ percent. These groups represent the tail of the skewed long distance expenditure distribution and should be evaluated further with a larger sample size in order to determine if the socioeconomic and demographic variables are capable of further explaining the variance in these customer segments. Until this is completed, an ultimate judgement on the efficacy of these variables in explaining long distance expenditure can not be made. However, in the final analysis, it appears that approximately 30 percent of the variance can be explained by these predictor variables, which is in line with the initial expectations when the project was undertaken.

An Approach Toward Empirical Model Building

Comparing linear regression and AID analysis, it is difficult to say which technique is better. Each offers certain advantages that the other does not, and each has inherent problems. AID is much more flexible in data analysis and data handling because it requires the smallest set of assumptions. At the same time, it is extremely compatible with the managerial viewpoint of the market place. Therefore, it has the advantage of better communicating the market research results. Finally, the technique brings into bold relief the relationships among the variables, which leads the researcher to think of the interactive effects of a set of predictor variables. This is very likely to broaden his inductive theorizing process. On the other hand, AID is lacking in inferential capability. It is much easier to display the data than to build predictive empirical models with AID. This is because AID is largely based on analysis of variance principles and therefore requires prior experimental or matrix designs to enable any inferences to be drawn from the analysis. A second disadvantage with AID is less parsimony in data analysis and model building; considerable computation and search is inherent in the technique and the branching process often is fairly complex and lengthy, which reduces its usefulness from a pragmatic control standpoint. Finally, as was demonstrated in this paper, AID requires large sets of observations which must be fairly well-behaved in their distributions over the predictor and the criterion variables. In other words, skewness presents a serious problem in AID analysis.

Linear regression is very powerful in developing parametric models, and provides a mechanism for establishing point and interval estimates for predictive purposes. On the other hand, it presumes the data to be linear, error free and additively related to the criterion variable.

In view of the fact that in a regulated industry there is a need to build powerful predictive models, a systematic approach is necessary to develop inductive models of telephone behavior based on large scale empirical data. The data analysis reported in this paper, together with the following procedural steps are recommended for inductive model building.⁸

1. Given a large scale data base, perform an initial AID analysis with as many predictor variables as are available or can be handled by the computer. The AID analysis will bring into bold relief the nature of the relationships among the variables resulting from a minimum set of restrictions with respect to the sample size of subgroups, split reducibility criterion and the priority ordering and coding aspects of the predictor variables. In short, a free AID analysis is recommended in this initial phase.
2. From the initial AID analysis, predictor variables should be selected for future analysis based on their explanatory power. The predictor variables should then be factor analyzed to estimate the degree of intercorrelations among them.
3. Choose a set of orthogonal predictor variables from the factor analysis results selecting the variable with the highest factor loading. The problems of error in measurement should also, be considered. For example it will generally be more advantageous to choose the age of the husband rather than the wife if both are loaded equally on a factor because of the possibility of response

error in the latter variable. Similarly, education would be preferred over income. At the same time, the researcher must watch for the possibility of creating an index variable, especially when several predictor variables contribute to an equal, but smaller, extent toward the eigenvalue of the factor. Such an index, by definition, would be a linear additive index.

4. Utilizing the selected orthogonal predictor variables, the researcher should perform a monotonic AID analysis. The restriction of monotonicity is more appropriate for managerial decision making since it will enable the researcher to develop models of a set of predictor variables which are split above or below a certain level.
5. Based on the monotonic AID analysis the predictor variables should be defined in terms of broad categories where a split occurred. For income in our data, this is likely to be below \$10,000, between \$10,000 and \$15,000, and above \$15,000. In the same way, a set of interactive predictor variables must be defined; for example, income above \$15,000 and eight or more rooms.
6. If the interactive effects are not substantial, as evidenced by the monotonic AID analysis, the simplest procedure would be to create a successive interval scale for each predictor variable based on AID categorization. This may result in a dichotomous scale or a discrete interval scale.

At this stage, a discriminant or regression model should be built in which the redefined variables developed from the prior analysis are the predictors and the phenomenon under investigation is the criterion variable. If the criterion phenomenon is dichotomous or classifactory, the discriminant model will be appropriate; however, a regression model should be used if the criterion variable is continuous and well behaved. The regression or discriminant model will then estimate a set of optimal weights for predictive and inferential purposes.

It is, however, possible that the interest is in building a model which takes into account each category of a predictor variable separately. This is possible by converting the regression or discriminant problem to a dummy variate analysis problem.

7. If there are strong interactions among the orthogonal predictor variables as evidenced from the monotonic AID analysis, it will be necessary to develop index variables based on the pattern of interactions. This should be relatively easy in view of the fact that the logical combinations are likely to be greatly reduced when the stage of performing a monotonic AID analysis is reached. The predictive model can be built from these index variables utilizing regression or discriminant analysis.

To summarize, several conclusions can be drawn from these efforts at inductive model building based on large scale data banks. First, it is extremely important to examine the quality of the data and the nature of the relationships among the variables. Without this critical examination, the researcher is likely to fall prey to a statistical or mathematical model popular at the time. Most of the recent model building in marketing has been based on management science techniques

which clearly attests to this problem. Second, it is very unlikely that a single statistical model such as stepwise regression, AID or discriminant analysis will be sufficient. The authors strongly suggest that a variety of statistical tools are sequentially necessary at various stages of inductive model building. Finally, it is unlikely that demographic factors alone will enable the researcher to build highly predictive models. The demographic factors, however, seem highly useful in segmenting the total population into subpopulations which may be the logical independent marketing segments requiring separate models.

Footnotes

1. This study is part of ongoing empirical research on the telephone behavior of both residence and business customers of the Bell System and was prepared under the auspices of the Market Research Section of the American Telephone and Telegraph Company in New York.

The authors wish to express their appreciation to Mr. N. J. Mammana, Director of Marketing Research for his support of the study and to Welling Howell who prepared and assembled the data and performed the computer analysis.

2. A Marvin Roscoe, Jr. is a Marketing Supervisor at A.T. & T. where he is responsible for developing analytic methodology for the Market Research Information System. Previously he was with the Bell Telephone Company of Pennsylvania and the Long Lines Department of A.T. & T. in various sales and marketing positions. He has a B.S.E.E. from Rensselaer Polytechnic Institute and a M.B.A. from the University of Pittsburgh.

Jagdish N. Sheth is presently Professor of Business and Research Professor at the University of Illinois. Prior to that, he was on the faculty of Columbia University and M.I.T. He received his Ph.D. at the University of Pittsburgh. He has also been visiting professor at the Indian Institute of Management, Calcutta, and Visiting Lecturer at the International Marketing Institute, Harvard University. Dr. Sheth is coauthor (with John A. Howard) of The Theory of Buyer Behavior, author of How Advertising Works and is a frequent contributor to business and scientific journals, especially in the area of marketing.

3. The authors are well aware of the controversy with regard to the usefulness of demographic factors in predicting consumer behavior (Yankelovich, 1964; Frank, 1968; Bass, Tigert, & Londale, 1968). However, given the regulated nature of the utility industry, it is necessary to understand and to be able to predict the impact of corporate strategies on different socio-economic segments of the population.
4. The term expenditure properly connotes the aspects of consumer buying behavior. Increasing the consumer's level of long distance expenditure is a very important consideration to the telephone industry. Since the distribution channel is always available and there are frequent periods of available capacity, increased calling during these periods can have very obvious economic implications to society.

5. There are several other methods for handling the multicollinearity problem, such as examination of the simple correlations or factor structure of the correlation matrix. In fact, due to the order bias built into stepwise regression, other methods should be used to reduce the collinearity problem.
6. Regression theory can handle nonlinear and interactive relationships, but these need to be developed a priori based on some theory of judgement. Without theory to suggest rational approaches, the number of nonlinear transformations and the combination of interactions are too many for regression analysis to solve efficiently.
7. By setting a low F value (0.01) all of the 14 demographic variables were permitted to enter in the final step of the regression analysis. The predictive power increased to 14.25 percent. These results were replicated using the UCLA BMD 03R Multiple Regression Program. Unfortunately, the additional explanatory power has the inherent problems of instability and multicollinearity.
8. The procedure is somewhat different from the two-stage AID-MCA linkage suggested by Sonquist (1970).

References

- Armstrong, J. S. & Andress, J. G. Exploratory analysis of marketing data: trees vs. regression. Journal of Marketing Research, Nov., 1970, VII, 487-492.
- Assael, H. Segmenting markets by group purchasing behavior: an application of the AID technique. Journal of Marketing Research, May, 1970, VII, 153-158.
- Blalock, H. M. Jr. Correlated independent variables: the problem of multicollinearity. Social Forces, 1963, 42, 374-380.
- Bass, F. M., Tigert, D. J., & Lonsdale, R. T. Market segmentation: group versus individual behavior. Journal of Marketing Research, Aug., 1968, V, 264-270.
- Carman, J. M. Multiple classification analysis without assumption of interval measurement, linearity, or additivity: a comparison of techniques. Proceedings of the Social Statistics Section, American Statistical Association, Dec., 1967, 260-270.
- Dixon, J. W. (Ed). BMD Biomedical Computer Programs, Los Angeles: University of California Press, 1971.
- Draper, N. R. & Smith, H. Applied Regression Analysis, New York: Wiley, 1966.

- Frank, R. E. Market segmentation research: findings and implications. In F. Bass et al. (Eds.), Application of the Sciences in Marketing Management New York: Wiley, 1968.
- Frank, R. E., Massey, W. F., & Wind, Y. Market Segmentation, Englewood Cliffs: Prentice Hall, 1972.
- Johnston, J. Econometric Methods, New York: McGraw-Hill, 1962.
- Lansing, J. B. & Kisk, L. Family life cycle as an independent variable. American Sociological Review, Oct., 1957, 22, 512-519.
- Lansing, J. B. & Morgan, J. N. Consumer finances over the life cycle. In L. Clark (Ed.), The Life Cycle and Consumer Behavior, New York: New York University Press, 1955, 36-51.
- Morgan, J. N. & Sonquist, J. A. Problems in the analysis of survey data and a proposal. Journal of the American Statistical Association, June, 1963, 58, 415-435.
- Sonquist, J. A. Multivariate Model Building, Survey Research Center, Ann Arbor: Institute for Social Research, University of Michigan, 1970.
- Sonquist, J. A. & Morgan, J. N. The Detection of Interaction Effects, Survey Research Center, Monograph No. 35, Ann Arbor: Institute for Social Research, University of Michigan, 1964.
- Staelin, R. A note on detection of interaction. Public Opinion Quarterly, Fall, 1970, 34, 408-411.
- Staelin, R. Another look at AID. Journal of Advertising Research, October, 1971, II No. 5, 23-28.
- U. S. Bureau of the Census. Methodology and Scores of Socioeconomic Status. Working Paper No. 15, Washington, D.C., 1963.
- Yankelovich, D. New criteria for market segmentation. Harvard Business Review, March-April, 1964, XLII, 83-90.

CONDOMINIUM DESIGN AND PRICING:
A CASE STUDY IN CONSUMER
TRADE-OFF ANALYSIS

John A. Fiedler
Associate Manager, Decision Systems Group
Market Facts, Inc.

Introduction

The concept of utility has been a familiar one to economists for over a century. Much of economics is based on a theory of buyer behavior in which the consumer is thought to allocate his resources so as to maximize his personal utility. In the late 19th century, it was fashionable among economists to attempt to assign numerical values to individuals' utilities for various quantities of goods. Finding those efforts lacking in predictive power, they have tended to abandon efforts to quantify utilities. However, the basic concept, the idea of the rational consumer, has remained.

It is possible that the consumer is unaware of the numerical values of his utilities, but that they may be revealed through his choices among product concepts which are varied in systematic ways.

This study is a practical application of the economists' traditional theory of buyer behavior. The study was based on a simple model which assumes:

Consumers can supply their rank orders of preference for various combinations of attributes that characterize the features of a product or service they intend to purchase.

Computational techniques recently available allow us to solve for a set of numbers for each consumer which adequately reproduce his rank orders of preference and which appear, therefore, to have the properties of utilities.

Finally, a consumer's choice from among several products or services can be predicted by combining his utilities for the attributes which characterize each good and determining which have the greatest utility for that individual.

This study was a cooperative project of Market Facts and a major builder. Because of this joint investment in the data and its analysis, some of the findings can be discussed; however, much of the data has been rescaled to protect the client's interests.

The Problem

The builder is currently developing a major residential complex on a unique 38 acre site, across the Hudson River from midtown Manhattan. When completed, the development will contain six high-rise buildings containing about 4000 condominium homes along with extensive recreational facilities.

Each thirty-one story building will consist of six different types of units. These are described in Figure 1 together with the range of prices for each type of unit in the first building. The price of a unit depends not only on its size, but also on its height and the direction of the view.

Figure 1
Price Schedule for Building One

Prices

Prices for the same type apartment homes vary because of floor, location and exposure.

Plan A. 3-bedroom, 3-bath corner apartment home \$64,500 \$78,700 <small>lobby floor 31st floor</small>	Plan D. 2-bedroom, 2-bath apartment home \$46,000 \$56,050 <small>lobby floor 31st floor</small>
Plan B. 2-bedroom, 2-bath deluxe corner apartment home \$56,500 \$68,000 <small>lobby floor 31st floor</small>	Plan E. 2-bedroom, 1½-bath convertible apartment home \$39,500 \$49,600 <small>lobby floor 31st floor</small>
Plan C. 2-bedroom, 2-bath deluxe apartment home \$52,000 \$65,200 <small>lobby floor 31st floor</small>	Plan F. 1-bedroom apartment home \$35,000 \$40,100 <small>lobby floor 31st floor</small>

The view of New York City across the river to the east contrasts sharply with the undramatic view of New Jersey to the west.

The floors above the first are identical with twenty apartments on each floor as shown in Figure 2.

The 614 units in the first building were priced in accordance with the developer's experience on other condominium projects. The larger the unit, the higher the unit, the better its view, the more expensive is its price.

The results of the first few weeks of sales for the first building are shown below:

Figure 2
Typical Floor in Building One

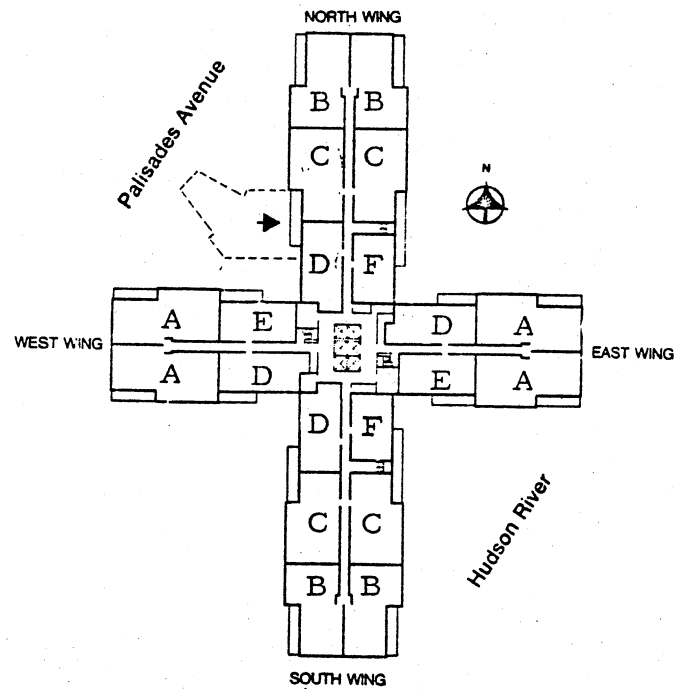


Table 1

Sales in Building One as of January 31, 1972

<u>Units</u>	<u>Number of Units</u>	
	<u>Sold</u>	<u>Available</u>
3-Bedroom (A)	13	124
2-Bedroom (B)	19	122
2-Bedroom (C)	25	122
2-Bedroom (D)	12	123
2-Bedroom (E)	19	62
1-Bedroom (F)	55	61
River View	121	338
No River View	22	276
TOTAL	<u>143</u>	<u>614</u>

Of the 276 units without a view only 22 or about 8 percent had been sold. Of the 338 units with a view, 121 or 36 percent had been sold. Some units were only selling on the lower floors, others only on the upper floors. The one-bedroom units were virtually sold out, while most of the three-bedroom units remained unsold.

Management realized it had pricing problems and adjusted the floor and view premiums in an attempt to create more realistic prices. However, the sales during the following months continued to be uneven. Cumulative sales through the end of July are shown in Table 2.

Table 2

Sales in Building One as of July 31, 1972

<u>Units</u>	<u>Number of Units</u>	
	<u>Sold</u>	<u>Available</u>
3-Bedroom (A)	43	124
2-Bedroom (B)	71	122
2-Bedroom (C)	75	122
2-Bedroom (D)	53	123
2-Bedroom (E)	51	62
1-Bedroom (F)	61	61
River View	248	338
No River View	106	276
TOTAL	<u>354</u>	<u>614</u>

At that time, almost 60 percent of the building was sold, a figure which would have been most acceptable to management if the sales had been evenly distributed throughout the building. But they were not. The two smallest units, E and F, were virtually sold out, while others remained unwanted. Seventy percent of the sales were units with a view across the river. It seemed as though the building might tip into the river.

The builder had already begun plans for a second building in the complex and needed a pricing formula which would provide a more even sellout of the building.

The Study

Accordingly, a study was undertaken which would utilize the model outlined earlier. The plan of the study was to:

Locate previous and prospective buyers of various-sized units in the development.

Collect from these respondents their rank orders of preference for condominium apartments having differing combinations of view, floor height, unit size, and price.

Calculate each respondent's theoretical utilities for the various levels of each attribute characterizing a condominium.

Construct a mathematical model to simulate the sell-out of a specified building under varying pricing formulae.

Find a set of prices which would cause the building to sell out as evenly as possible.

Questionnaire Design and Field Survey

It was decided that any apartment in the building could be characterized by a combination of the relatively few levels for the four variables to be measured. Further, certain levels were not applicable to all respondents and a respondent interested in a unit of a specified size was questioned only on those which were relevant to him. The levels of the attributes are shown in Table 3. Each of the four "floors" studied actually represented about eight floors in the building.

Approximately two hundred respondents were selected from among visitors to the model units of the development site and from among previous buyers returning to the development office to select various decorating options.

Table 3

Levels of Attributes Measured in Survey

<u>Attribute</u>	<u>Levels</u>				
Floor --	28th	20th	12th	4th	
View --	View of River		No View of River		
Purchase Price --	\$46,000	\$52,000	\$59,000	\$66,000	\$74,000
	\$49,000	\$55,000	\$64,000	\$73,000	\$82,000
Unit Type --	Plan A		Plan C	Plan E	
	Plan B		Plan D	Plan F	

The tasks given the respondent were simple. A sample page from the questionnaire is shown below.

Here the respondent is asked to imagine eight possible apartments, each differing only in floor and view. If he could have any of the eight, which

would be his first choice? Most respondents placed a 1 in the upper-left box. Now if that unit were unavailable, which would be his choice? This procedure was repeated until the respondent had provided his rank orders of preference for all eight units. Using such a ranking procedure, each of the four attributes was compared to each other.

Figure 3
Sample Questionnaire Page

You could have an apartment
with a view . . .

And could be on the ...	Toward the Hudson River	Away from the Hudson River
28th Floor		
20th Floor		
12th Floor		
4th Floor		

In addition to the trade-off data just described, standard demographic and classification data were collected.

By merely tabulating the survey results the client was provided with a type of data unique to the housing industry which provided answers to questions such as:

Would people rather live in a large unit with no view or a small one with a view?

What proportion of people interested in one-bedroom apartments would be willing to pay an extra \$25 a month for a second bedroom?

Would people be willing to move from the twenty-eighth floor to the twelfth in order to keep the view across the river?

Computation of Utilities

Direct examination of the trade-off data, however, allows only two attributes to be compared at a time. Since each apartment unit is characterized by four attributes it is desirable to compute "utilities" for each level of each attribute so that these may be combined to predict each respondent's choice from among various types of apartments.

Figure 4
Sample Questionnaire Data

		You could have an apartment with a view . . .	
		Toward the Hudson River	Away from the Hudson River
And could be on the ...			
28th Floor		1	4
20th Floor		2	5
12th Floor		3	7
4th Floor		6	8

The computational procedure used is similar to pairwise nonmetric factor analysis and has been described in detail in a paper by Johnson (1972). A short example will suffice to explain the technique here. Suppose a respondent has given us ranked data as shown in Figure 4. The procedure solves for a number for each floor and one for each of the two types of views. These numbers are determined so that their products have the same (or nearly the same) rank orders as the original data. An example is given below.

Figure 5

PAIRWISE PRODUCTS OF UTILITIES

		View	No View
		.7	.3
28th Floor	.4	.28 (1)	.12 (4)
20th Floor	.3	.21 (2)	.09 (5)
12th Floor	.2	.14 (3)	.06 (7)
4th Floor	.1	.07 (6)	.03 (8)

It can be readily seen that these numbers have the same rank order as the original data. This is not always the case. When an attribute is compared to several others, the respondent may be inconsistent in his preferences so that no set of numbers can be found which will fit the data perfectly.

We required a measure of how well the utilities fit the data. An appropriate statistic is Kendall's tau, which involves a count of the pairs of ranks which are in the right order and those which are in the wrong order. Tau is the difference between two such proportions or

$$\tau = \frac{\text{number of rights} - \text{number of wrongs}}{\text{number of rights} + \text{number of wrongs}}$$

In the case of the above example, tau has a value of 1.0. A tau of 0.0 would indicate no order relationship between the predicted value and the data.

Figure 6-11 show the six trade-off matrices supplied by an actual respondent. This respondent happened to be a married man with higher than average income with no children who had been looking for a new home for about three months.

In the first matrix, he tells us his preference for the river view totally dominates his preference for floor height. (See Figure 6)

Figure 6

First Trade-Off Matrix: Floor vs. View

You could have an apartment with a view . . .

	Toward the Hudson River	Away from the Hudson River
And could be on the . . .		
28th Floor	1	5
20th Floor	2	6
12th Floor	3	7
4th Floor	4	8

In the second matrix we see that although he was willing to give up in floor height to keep that important view, he is not willing to pay an extra \$60 a month to have it. (See Figure 7)

Figure 7

Second Trade-Off Matrix: View vs. Price

You could get an apartment at this price . . .

	\$52,000 (\$440 per month*)	\$59,000 (\$500 per month*)	\$66,000 (\$560 per month*)	\$74,000 (\$625 per month*)
And have a view . . .				
Toward the Hudson River	1	3	5	7
Away from the Hudson River	2	4	6	8

*Monthly cost includes all payments. Taxes, maintenance, principal and interest on a 30-year, 7-1/2% mortgage with 20% down.

From the next matrix we learn he would prefer to give up the river view rather than to live in the small two-bedroom apartment with a view.

Figure 8

Third Trade-Off Matrix: View vs. Unit Type.

You could have this apartment . . .

	Plan B (2-bedroom, 2-bath de- luxe corner apartment)	Plan C (2-bedroom, 2-bath de- luxe apartment)	Plan D (2-bedroom, 2-bath apartment)	Plan E (2-bedroom, 1-1/2 bath convertible apartment)
And could have a view . . .				
Toward the Hudson River	1	2	3	7
Away from the Hudson River	4	5	6	8

In this matrix, unit price is traded off against unit type. Again there is a total rejection of Plan E. Plan C at \$52,000 is preferred to Plan D at \$52,000 and so on.

Figure 9

Fourth Trade-Off Matrix: Price vs. Unit Type

You could have this apartment . . .

	Plan B (2-bedroom, 2-bath de- luxe corner apartment)	Plan C (2-bedroom, 2-bath de- luxe apartment)	Plan D (2-bedroom, 2-bath apartment)	Plan E (2-bedroom, 1-1/2 bath convertible apartment)
And pay this price . . .				
\$52,000 (\$440 per month*)	1	4	9	13
\$59,000 (\$500 per month*)	2	5	10	14
\$66,000 (\$560 per month*)	3	7	11	15
\$74,000 (\$625 per month*)	6	8	12	16

The next ranking sheet shows that this respondent preferred Plan B on any floor except the 4th. If that were not available he would prefer Plan C, again ruling out the 4th floor. He does prefer Plans B and C on the 4th floor to either Plan D or Plan E on any floor.

You could have this apartment . . .

And could be on the . . .	Plan B (2-bedroom, 2-bath de-luxe corner apartment)	Plan C (2-bedroom, 2-bath de-luxe apartment)	Plan D (2-bedroom, 2-bath apartment)	Plan E (2-bedroom, 1-1/2 bath convertible apartment)
28th Floor	1	4	9	13
20th Floor	2	5	10	14
12th Floor	3	7	11	15
4th Floor	6	8	12	16

In the last matrix, we see he would prefer to spend \$59,000 and live on the 28th of 20th floors rather than to spend \$52,000 and live on the 4th floor.

You could get an apartment at this price . . .

And could be on the . . .	\$52,000 (\$440 per month*)	\$59,000 (\$500 per month*)	\$66,000 (\$560 per month*)	\$74,000 (\$625 per month*)
28th Floor	1	4	9	13
20th Floor	2	5	10	14
12th Floor	3	7	11	15
4th Floor	6	8	12	16

These data were supplied to the utility calculating program with the results as shown in Table 4.

Table 4

Example of a Respondent's Utilities

<u>Attribute: Level</u>	<u>Utility</u>	<u>Attribute: Level</u>	<u>Utility</u>
Floor: 28th	.315	Price: \$52,000	.738
20th	.311	\$59,000	.217
12th	.271	\$66,000	.035
4th	.103	\$74,000	.010
River View	.769	Unit: Plan B	.471
No View	.231	Plan C	.403
		Plan D	.125
		Plan E	.001

$$\text{tau} = \frac{441 - 3}{444} = \frac{438}{444} = .986$$

When these utilities are cross-multiplied and their products rank ordered, we find that the respondent's data were correctly predicted for four of the six matrices. There were three pairwise errors of prediction in the remaining two matrices. The calculation of the resulting tau of .986 is shown.

How typical was this respondent of the sample? Table 5 shows the distribution of tau for the sample.

Table 5

Distribution of Tau

<u>Tau</u>	<u>Respondents</u>	
	<u>No.</u>	<u>%</u>
1.000	25	13.3
.950 - .999	51	27.1
.900 - .949	56	29.8
.850 - .899	27	14.4
.800 - .849	17	9.0
.400 - .799	12	6.4

An arbitrary cutoff was made, and ten respondents whose taus were less than .775 were eliminated from further analysis. Our sample respondent was at the 79th percentile.

Modeling the Building Sellout

The first way the utilities were employed was to test the model by a simulation to "predict" the sellout of the first building. For the purpose of the model, we viewed the building as consisting of the six types of units, with or without a view, on one of four floor levels. Thus there were 48 different combinations of apartment configurations ($6 \times 2 \times 4 = 48$).

Appropriate prices for each of these units were obtained using the client's pricing schedule. Each respondent's utility for each apartment configuration was then computed by multiplying together his utilities for the levels of the attributes comprising the particular unit. An example, using the utilities for the respondent seen earlier, shows his utilities for eight of the possible 48 units.

Table 6Respondent's Utilities for Selected Apartments

<u>Apartment Configuration</u>	<u>Utilities</u>	<u>Overall Utility</u>
1) B Unit, view, 12th floor, \$66,000	.471 x .769 x .271 x .035 =	.0034
2) B Unit, no view, 20th floor, \$59,000	.471 x .231 x .311 x .217 =	.0073
3) C Unit, view, 4th floor, \$59,000	.403 x .769 x .103 x .217 =	.0069
4) C Unit, no view, 28th floor, \$59,000	.403 x .231 x .315 x .217 =	.0064
5) D Unit, view, 28th floor, \$52,000	.125 x .769 x .315 x .738 =	.0223
6) D Unit, no view, 20th floor, \$52,000	.125 x .231 x .311 x .738 =	.0066
7) E Unit, view, 12th floor, \$52,000	.001 x .769 x .271 x .738 =	.0002
8) E Unit, no view, 20th floor, \$52,000	.001 x .231 x .311 x .738 =	.0001

Note that the first unit shown, which would have been the most attractive of these if cost were not an issue, has a relatively low overall utility. The apartment with the greatest appeal is the fifth unit shown, combining the lowest price with a view and having an acceptable floor plan on a floor above the fourth.

In most cases, the price levels from which utilities were calculated did not correspond exactly to actual market prices for the units. In these cases, interpolation was used to estimate a respondent's utilities for market prices.

A respondent would theoretically be most likely to purchase that type of unit for which his utility is the highest. However, that unit might be sold out when he is ready to buy; in such a case, we would expect him to purchase that apartment for which his utility was the second highest. However, if all but a few types of units had sold out and those remaining had very low utilities for a particular respondent, it seems quite likely that he would buy elsewhere rather than purchase such a unit. It was therefore necessary to estimate how far down in his choice hierarchy of possible units an individual would be likely to make a purchase. Various possibilities were computed and compared against the actual sellout of the first building. While solutions were not particularly sensitive to the value chosen, the assumption that an individual would be willing to purchase any of the three units for which he had the highest utility did provide the best fit to the data.

The marginal totals are shown below.

Table 7

Actual vs. Predicted Sellout of Building One

<u>Units</u>	<u>Sellout</u>	
	<u>Predicted</u>	<u>Actual</u> (7-31-72)
3-Bedroom (A)	42	43
2-Bedroom (B, C, D, E)	253	250
1-Bedroom (F)	59	61
View	248	248
No View	106	106
28th Floor	53	
20th Floor	85	
12th Floor	110	
4th Floor	106	

Although data by floor were not available, the client confirmed the fact that more units had been sold on the lower floors. While the model performed satisfactorily in predicting the marginal totals, as shown, it did less well in predicting the sellout of units with specified sizes, prices and views. The client attributed much of this error to the changes that were made in pricing schedule during the first seven months of sales. It was therefore decided to accept the model and proceed to explore sellout rates as a function of pricing schedules for the second building.

Determining an Optimal Price Schedule for Building Two

The second building's configuration is quite different from that of the first. Because of its location on the site, fewer units will have a view across the river. There will also be fewer three-bedroom (A) units, and more of the larger two-bedroom units (B, C, D). The next step of the analysis was to assume an initial set of prices for the different types of units in the new building and input prices to the model. The outcome of this simulation is shown in Table 8.

This initial pricing schedule did even a poorer job than Table 8 indicates. When the results of the simulation were examined in greater detail, comparing the predicted sellout with the target, we found errors such as: the model predicted the sale of 41 B units on the "28th floor" without a view, but only 15 were available; the model predicted the sale of only 11 units on the same "28th floor" with a view out of a total 31 available. A useful measure of the error of prediction is the root mean square, a weighted average error. For the forty-four types of units in the simulation, the root mean square error was 12.7 units.

A number of judgmental attempts were made to create a better pricing schedule which would reduce this error and provide a more even sellout. These were not only time-consuming, but mostly unproductive as well.

Table 8Predicted Sellout of Building Two

<u>Units</u>	<u>Predicted by Model</u>	<u>Target</u>	<u>Error</u>
28th Floor	134	169	-35
20th Floor	192	169	+23
12th Floor	182	169	+13
4th Floor	170	169	+ 1
View	326	310	+16
No View	352	366	-14
Floor Plan A	51	62	-11
Floor Plan B	209	184	+25
Floor Plan C	115	184	-69
Floor Plan D	132	123	+ 9
Floor Plan E	29	31	- 2
Floor Plan F	142	92	+50

Accordingly, we decided to let the model search for a set of prices that would theoretically sell the building out evenly. An iterative procedure was used that simply increased slightly the prices of those units where predicted sales exceeded the number available and decreased the prices of those units which were not predicted to sell well enough.

The model did not find a set of prices that performed perfectly, though it did converge on an adequate solution. The results are shown in Table 9.

For the detailed list of 44 types of units, the root mean square is 3.0. This is about 75 percent reduction in error, and appears to be as low a value as is possible with the limited sample size and the integer nature of the predictions.

All that remained was to formalize these prices through the creation of a corresponding formula to price all 676 units in the new building. This was done by solving for a set of base prices for each unit, and premium adjustments for floor and view which corresponded most closely to the theoretically optimal prices.

Conclusion

The only real test of the model will be how evenly the new building does sell out. Our predictions are on record.

In addition to pricing a building, the data can be used as an input to planning the configuration of future building. We should theoretically be able to determine whether certain units should be built only with or without a view, or only on certain floors.

Table 9Final Predicted Sellout of Building Two

<u>Units</u>	<u>Predicted by Model</u>	<u>Target</u>	<u>Error</u>
28th Floor	160	169	- 9
20th Floor	177	169	+ 8
12th Floor	166	169	- 3
4th Floor	175	169	+ 6
View	306	310	- 4
No View	372	366	+ 6
Floor Plan A	42	62	-20
Floor Plan B	187	184	+ 3
Floor Plan C	176	184	- 8
Floor Plan D	117	123	- 6
Floor Plan E	42	31	+11
Floor Plan F	114	92	+22

The applications of this model are obviously not limited to housing. The model has also been used to study intercity air travel, the market for sophisticated office equipment, the operation of urban mass transit systems, as well as in the areas of financial services and government regulation.

Reference

Johnson, Richard M. Trade-Off Analysis: A Method for Quantifying Consumer Values. Market Facts Inc., Chicago, 1972.

FORECASTING DEMAND FOR A
NEW MODE OF TRANSPORTATION

J. D. Davidson
Senior Operational Research Designer
Air Canada

Introduction

Starting in mid-73, a subsidiary of Air Canada will be operating a STOL (short take-off and landing) service between downtown Montreal and downtown Ottawa. The service is an experiment sponsored by the Canadian Government using a Canadian built aircraft--the Twin Otter. The O.R. group was asked to help in the marketing of STOL by developing a model for forecasting traffic and predicting the effect of changing the levels of service.

This experiment will offer a fifth mode of travel between the two cities: the existing modes are car, bus, train and CTOL (Conventional take-off and landing). It is difficult to use historical data from the four existing modes to predict traffic on STOL even if it existed because of the fact that STOL will offer a service outside existing experience. In particular, it will offer a shorter door-to-door time than any existing mode and will probably cost more.

-Data was gathered by a consumer research company, Market Facts of Canada Ltd, which was commissioned to make a random telephone survey of 21,000 households followed up by 1,055 selected personal interviews. After we had started work on model design we found that Market Facts' parent company in Chicago had already developed two models which would meet our requirements very nicely. We therefore made use of these two models. The method uses attitudinal or "soft" data on such things as meal service and comfort as well as "hard" data on such things as time and price.

One of the major constraints was that the model had to be portable i.e. STOL could be sold as a package (aircraft, avionics, operating and marketing know-how) anywhere in the world. In particular, all the demographic and travel statistics which are available in Canada could not be used as they might not be available for another city pair elsewhere. They were, of course, used as a cross-check, sometimes to our chagrin. The acceptability of the sample is discussed later.

The project was carried out in five stages:

1. Qualitative survey to find out why travellers choose or reject each mode of travel and to judge their reaction to the proposed STOL service.
2. Model design--to build a model which will incorporate attitudinal data and predict the response to the new STOL service.
3. Data gathering by means of a large number of random telephone calls to households in Montreal and Ottawa followed up by personal interviews of a selected sample.
4. Model testing and calibration--tuning the model so that it can reproduce the existing modal split.

5. Forecasting and evaluation--adding STOL to the existing four modes and evaluating the effect of different marketing strategies on the STOL market share.

Research Surveys

We started with qualitative research to determine consumer attitudes to the proposed STOL service. This was done by group depth interviews. We then took the groups for a ride in the Twin Otter and followed up with further group interviews. The change in perceptions was remarkable. As a result of all the interviews we had a lot of qualitative information on how people perceived the advantages and disadvantages of each of five modes of travel. We also gained some knowledge of which attributes are important to travellers and which are regarded as trivial.

The qualitative survey was then followed by a quantitative survey--a random telephone survey of 15,000 households in Montreal and 6,000 in Ottawa. The number of phone calls was chosen to give a statistically acceptable sample of travellers in each city. That gave us demographic and travel information: number of trips made in the last 12 months, mode used on each trip and reason for the trip (business or non-business). The only other data we needed was the population over 18 of the two cities: Ottawa 370,000; Montreal 1.8 million. One interesting fact which emerged: half the population of Ottawa has visited Montreal in the last 12 months. We cannot think of another city pair where half the population of one has visited the other in a 12 month period.

The Interview Stage

As a result of the qualitative research, it was found that all five modes could be described in terms of 13 independent attributes. Each attribute was defined in terms of levels e.g. the attribute "time" had three levels-- 1 hour, 2 hours and 3 hours. Some attributes had two levels and some had four but the average was three. Hence we had 39 points: 13 attributes X 3 levels = 39.

Table of Attributes Used to Define
Modes of Travel

Attribute	Level 1	Level 2	Level 3	Level 4
1 Time (door-to-door)	3 hours	2 hours	1 hour	
2 Cost (door-to-door)	\$32	\$16	\$8	\$4
3 Schedule (departures/day)	5	16	22	anytime
4 Meal Service	nothing	coffee, soft drinks	coffee, soft drinks, bar	full meal and bar
5 Comfort	narrow seat	medium seat	wide seat	

The purpose of the interview was to obtain relative utilities for the 39 points. Used in this sense, a utility is simply a number between 0 and 1. It is further assumed that the pairwise products of utilities reflect a person's trade-off preferences. The interview, which took over an hour, required each respondent to fill in 21 matrices, ranking his preference for pairs of attributes. Suppose we knew what a person's utilities for time and cost were. We could then obtain utilities for pair combinations by multiplication. For example: $U(1 \text{ hour}) \times U(\$8) = .7 \times .25 = .175$. This assumes that the attributes are independent, something we tried to ensure when deriving them.

Matrix for Pairwise
Comparison of Attributes

Attribute	Time	1 hour	2 hours	3 hours
Cost	Utilities	.7	.2	.1
\$4	.6	.42 1	.12 3	.06 5
\$8	.25	.175 2	.05 6	.025 8
\$16	.1	.07 4	.02 9	.01 11
\$32	.05	.035 7	.01 10	.005 12

Having calculated utilities for each of the twelve combinations in the above matrix we could then rank them from 1 to 12 according to the rank order of the utilities.

In practice, of course, we can't expect the respondent to tell us what his utilities are but it is a simple task for him to rank order the cells in the above matrix according to his preferences. He was asked to do so without thinking of any particular mode. Once we have this rank ordered preferences we can infer his utilities using a technique known as pairwise non-metric multidimensional scaling. The particular computational technique which was used is similar to techniques for non-metric multidimensional scaling which have been developed by Kruskal (1964a-, 1964b-, 1968) and Johnson (to be published). A recent application may be found in Gibson et al. (1972). Mathematically, the problem is to find utilities for 39 attribute levels such that the pairwise products of utilities have rank orders similar to those provided by the respondent. In the iterative procedure, an initial solution is tested and the utilities modified so that the pairwise products are more like the respondent's rank ordered preferences. The programme seeks to minimize a value called "badness of fit." Iterations will continue in an attempt to reduce badness of fit until:

- 1 the badness is reduced to an acceptable level
- or 2 the badness fails to improve significantly in 5 iterations
- or 3 a specified number of iterations is reached

The programme which converts pairwise rank ordered preferences to utilities was developed by Johnson from the procedure described by Kruskal (1968) and is the property of Market Facts Inc. This is the first of the two models. This is an example of conjoint measurement--a technique which allows us to estimate a consumer's value system by observing his behaviour. The way he trades off one attribute against another tells us about his values for both attributes jointly; it might not be possible to measure his values by taking the attributes one at a time. Twenty one attribute pairs were ranked--each attribute was compared with a least two others and important attributes like time and cost were compared with about five others.

The Forecasting Model

This is the second model. The 1,055 respondents were divided into classes: 4 geographical areas; business/non-business; light/heavy travellers; most often by air/most often by car/most often by other surface mode. This was done so that we could weigh each respondent correctly in order to relate respondent behaviour to the results of the telephone survey. Although the telephone survey was a random sample, the interviewees were not. The sample was structured to obtain equal representation in each class.

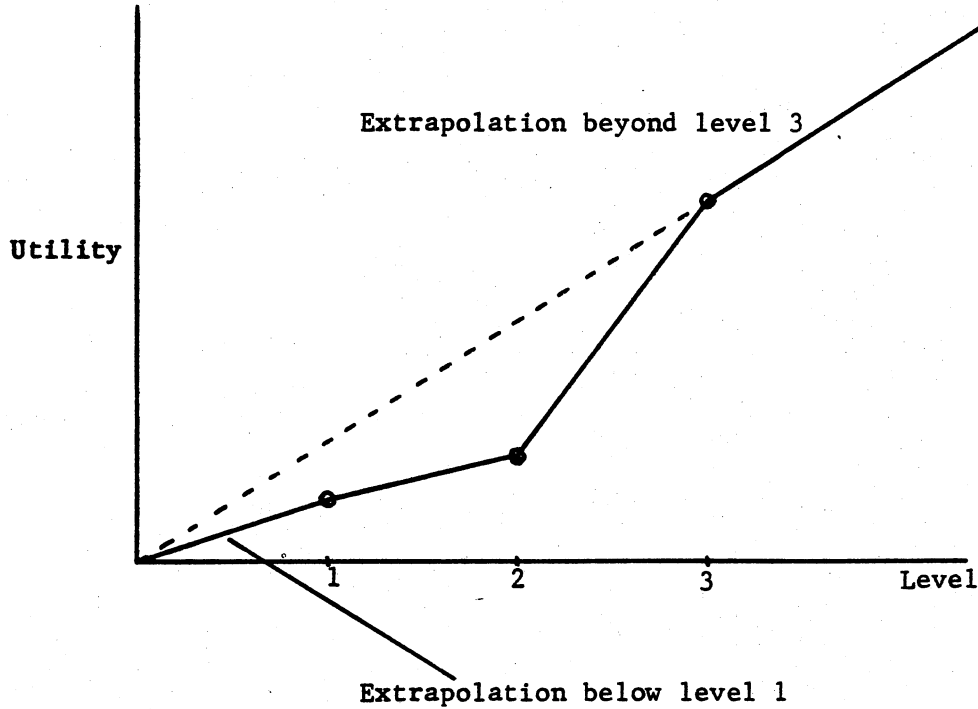
To calibrate the model, we had to input the attributes of the four existing modes and ensure that the model produced the correct modal split.

Definition of Existing Modes

Level	Attribute	Travel Mode			
		Car	Bus	Train	Air
1	Trip time is 3 hours				
2	Trip time is 2 hours	1.5	.7	.8	1.9
3	Trip time is 1 hour				
1	Cost is \$32				
2	Cost is \$16	4.0	3.8	3.2	1.8
3	Cost is \$8				
4	Cost is \$4				

Level 1.5 for the trip time by car implies that the time taken lies half way between level 1 and level 2 i.e. 2 1/2 hours. Level 4 for the cost of a car trip implies that the door-to-door cost of a one way trip is \$4 per head. Level of 1.9 for the trip time by air represents 2 hours 6 minutes.

This process was repeated for all 13 attributes. Although we had calculated respondents' utilities for levels 1, 2, 3, etc. for each attribute, it was possible to interpolate between these levels and even extrapolate beyond them. When designing the questionnaire, we had tried to choose a range for each attribute which would encompass all possibilities to avoid having to extrapolate. The model assumed straight line interpolation between points. The diagram shows how extrapolation was done.



A respondent's relative likelihood of taking each mode was calculated as follows:

$$U_{\text{Car}} = K_C (U_{C1} U_{C2} \dots U_{C13})^\alpha$$

$$U_{\text{Bus}} = K_B (U_{B1} U_{B2} \dots U_{B13})^\alpha$$

$$U_{\text{Train}} = K_T (U_{T1} U_{T2} \dots U_{T13})^\alpha$$

$$U_{\text{Air}} = K_A (U_{A1} U_{A2} \dots U_{A13})^\alpha$$

Here we have adopted the convention that U_{C1} represents the respondent's utility for the level of attribute 1 for car i.e. his utility for level 1.5 for time. The respondent's "utility" for each mode was calculated by multiplying together the 13 utilities for levels of each attribute. The constant K_i is known as the mode effect and is used to make a correction for any unexplained variable. We found that $K_C = K_B = K_T = 1$ but that the mode effect for air was considerably less than 1. The value of K_A was determined by trial and error and was required because otherwise the model consistently over-predicted the number

of air travellers. Only a fraction of those who would find air attractive actually took the plane. Perhaps the 13 attributes do not completely define each mode: we think there may be a 14th attribute. K_i was set at zero if a respondent answered no to the question "Would you ever travel to Ottawa by air? (bus? train? car?)."

The value of the exponent α was also determined empirically and came out at .5. There are two major reasons why we expected the exponent to be less than 1. Although we tried to devise 13 independent attributes there is inevitably some redundancy and the exponent makes a correction for things being counted more than once. The second reason is that the algorithm used to convert rank orders to utilities has a tendency to maximize the difference between the lowest level of each attribute and the highest. The exponent therefore corrects for the fact that a respondent's behaviour may operate within a narrower range than that represented by the utilities. Interestingly, in a previous study done by Market Facts Inc., there were 26 attributes and an exponent of .25--twice the number of our attributes and half the exponent. It should be noted that use of the exponent does not change the rank order of anything. A person's response to changes in levels is "damped" (in engineering terms) but rank orders are preserved.

Having calculated a person's "utilities" for each mode, the model then assumes that his trips are made on each mode in proportion to his utilities for each mode. A weighting is then applied to the trips to give the number of trips by persons in that class in the telephone survey. The total number of trips made is then calculated by multiplying by the ratio of population to telephone survey size.

Forecast

Having defined the four existing modes and calibrated the model, the predictions of modal split compared closely with the actual modal split. The actual modal split was obtained from the telephone survey by taking the most recent trip of each traveller. The comparison between actual modal split and the projection produced by the model is given below. These are not the real number (they have been changed to protect the innocent) but the discrepancies shown are of the same order.

We then inserted the fifth mode--STOL--and defined it in terms of the 13 attributes. The projected market share is shown in the table below. STOL was given the same exponent and mode effect as Air.

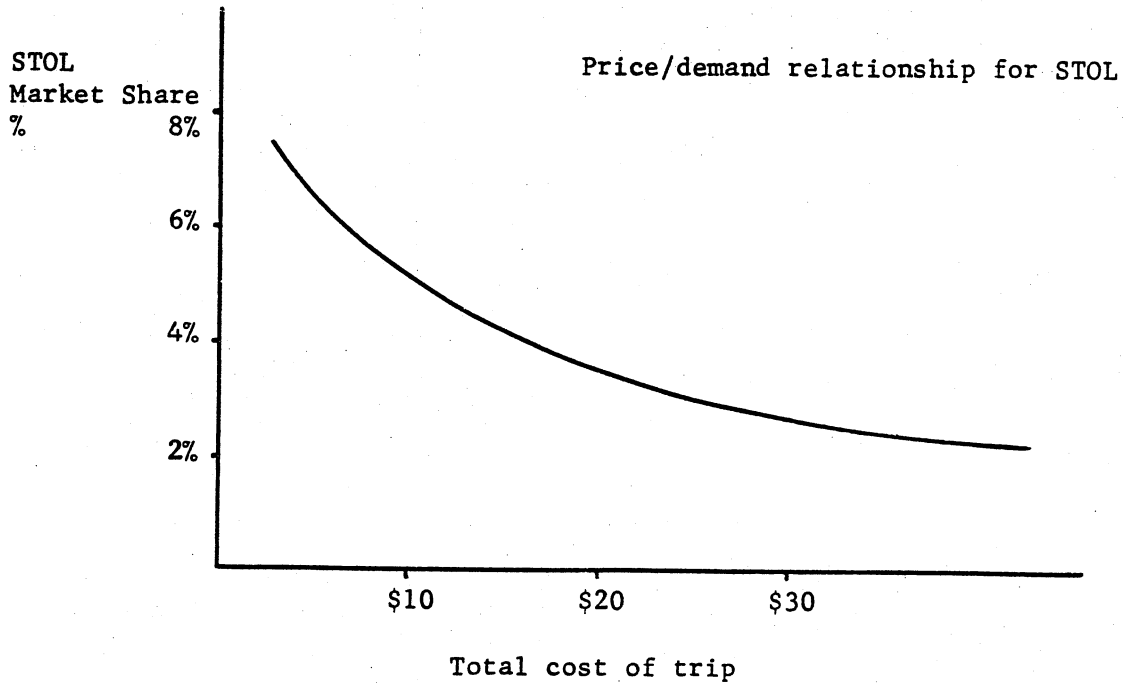
Several runs were made, changing the levels of the STOL attributes to determine the effect on traffic. We tried different fares, different trip times, beverage service etc. so that the STOL marketing manager could put together an optimum package and have a couple of tricks up his sleeve if things did not work out as expected.

Curves were drawn showing the change in traffic with change in levels of an attribute. The curve below shows the response of market share for STOL to changes in fare.

An additional benefit was the ability to identify who the most likely travellers on STOL would be. The relative utilities ($U_C, U_B, U_T, U_A, U_{STOL}$) were printed out for each respondent and the top 10 percent who had the

Protection of Market Shares

	Actual	<u>Original Projections</u>		<u>New Projections</u>		Change
	<u>Share</u>	<u>Shares</u>	<u>Trips</u>	<u>Shares</u>	<u>Trips</u>	<u>in share</u>
STOL	-	-	-	4.16	146,000	
Air	3.22%	3.22%	113,000	2.83	99,000	-12%
Train	10.94	10.39	364,000	9.94	348,000	- 4%
Bus	15.14	15.50	543,000	14.06	492,000	- 9%
Car	70.71	70.89	<u>2,481,000</u>	69.11	<u>2,419,000</u>	- 3%
			<u>3,500,000</u>		<u>3,500,000</u>	



highest utility for STOL could be examined individually. For promotional purposes, the Marketing Manager needs to know where they live, which language they speak, how much they earn, their purpose of travel, etc.

PROBLEM--OR WHAT WE WOULD DO DIFFERENTLY NEXT TIME

Let there be no doubt that the model has exceeded our expectations and has performed very well in a way that inspires confidence in the user. As in all projects, however, with the benefit of 20-20 hindsight we can see things that might be done a little differently. Progress comes with learning from our experiences.

One of our problems was in forecasting total travel between Montreal and Ottawa. There are statistics for the common carriers (which have to be taken with a large pinch of salt) but none for car which is twice as large as all the other modes put together. The requirement for portability meant that we should have been able to predict total traffic from the results of the telephone survey. We did not have too much faith in those but in the absence of any other data, they would serve the purpose. A cross check with data from other sources led us to believe that the total number of trips estimated from the telephone survey was too high and an examination of the data showed the reason.

Frequency of Travel Data from Telephone Survey

No. of trips made last year	1	2	3	4	5	6	7	8	9	10	11	12
No. of respondents	462	171	91	46	30	14	6	6	-	30	2	9
Trips p.a.	13	14	15	16	17	18	19	20	21	or more		
Respondents	-	1	7	-	-	-	-	-	9	15		

It seems that the appeal of nice, round numbers like 10, 12, 15 and 20 was too much for many people to resist. This is a market research problem. Is there some way we can improve the accuracy of people's recall? Is it realistic to expect people to be able to recall accurately when the number of trips made was more than about 4?

Perhaps a better way would be to ask a respondent about the trips he made in the last month and then forecast annual volume making allowance for seasonality. This implies that external data on seasonality is available and that might affect the portability of the model.

Another problem is that the interaction between interview task, interviewer and interviewee may have resulted in some misunderstanding. Consider the following utilities for trip cost:

Cost	-	\$32	\$16	\$8	\$4
Utility	-	.633	.286	.007	.073

This respondent was a French-speaking female aged between 18 and 24 with an income between \$5,000 and \$7,000 who went to Ottawa three times last year by car for non-business reasons. Her preferences seem to be inconsistent with her behaviour.

We would like to develop criteria to accept or reject a respondent based on the credibility of his responses. If a respondent tells us his preferences, what justification have we for rejecting him because it does not make sense to us? Consider this same respondent's utilities for time.

Trip time	-	3 hours	2 hours	1 hour
Utilities	-	.288	.008	.702

It may not look terribly rational but is not nonsense. Should it be rejected? Another approach might be to throw away the top 5 percent of respondents whose theta--the measure of badness of fit--is highest.

A considerable amount of work and soul-search went into determining the attributes and their levels and into the design of the questionnaire. The questionnaire was pretested, modified, pretested again and modified again before the field work of 1,055 interviews was conducted. Despite that, I believe that even more pretesting over a longer time would significantly improve the quality of the final results. The results of the field work indicate that some of the questions were not properly understood, some were not well defined (how do you define "ambiance"?) and others which were not meaningful could have been dropped. The choice of attributes and design of the questionnaire is a major task which should not be underrated.

Montreal and Ottawa were split into two zones each on the assumption that access and egress times to the airport and train and bus stations would vary by where the respondent lived. It turned out that a two-zone breakdown by city was not useful and that the breakdown should have been much finer--say 10 zones per city.

I suspect that the model may overestimate the number of travellers that will switch from car to STOL. The reasons is that we have made no allowance for groups. Car is the only mode where four people can travel as cheaply as one. An individual travelling on his own, for example, may choose between the car at \$4 and an air fare of \$14. If a party of four is travelling together then the choice is between a dollar a head or a total air fares bill of \$56. The model thus understates the attractiveness of car for people travelling in groups.

The Marketing Manager of STOL is using this model continuously and finding it invaluable in formulating his marketing plans for STOL. The acid test of how good the model actually is will only come in mid-73 when STOL goes into operation.

Perhaps it should be emphasized once again that this operation is an experiment, not a profit-making venture. With only 11 seats in it, (normally it has 20) the Twin Otter is not economic in this type of service. However, given the operating economics of a viable aeroplane such as the DHC-7, this model would prove to be very valuable in defining a profitable and attractive service. We anticipate that the approach taken in this study will have many more applications within Air Canada, not just for STOL.

References

- Kruskal, J. B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 1964a, 29, pp. 1-27.
- Kruskal, J. B. Nonmetric multidimensional scaling: a numerical method. *Psychometrika*, 1964b, 29, pp. 115-29.
- Kruskal, J. B. How to use M-D-SCAL, a program to do multidimensional scaling and multidimensional unfolding. Bell Telephone Laboratories, Murray Hill, N.J., March 1968 (mimeographed).
- Johnson, R. M. Pairwise nonmetric multidimensional scaling. *Psychometrika*, to be published.
- Gibson, R. E., Neidell, L. A., & Teach, R. D. Performance space analysis for an industrial product. *Operational Research Quarterly*, 1972, 23, pp. 125-38.

CONSUMER MENU PREFERENCE: AN APPLICATION
OF ADDITIVE CONJOINT MEASUREMENT

Paul E. Green
Yoram Wind
and
Arun K. Jain¹
University of Pennsylvania

In a recent JMR article Green and Rao (1971) presented an exposition of conjoint measurement procedures and commented on the fact that application of these techniques was still in its formative stages. This note describes one such application in the area of consumer preferences for menus. Specifically, additive conjoint measurement is applied to respondent evaluations of two-, three- and four-component restaurant menus in order to derive part-worth contributions of the components: appetizer, entree, dessert and price, to the total utility of a menu.

Two main questions guided the research design² and subsequent analysis:

1. Are utilities for menu components describable by an additive utility model, i.e., where total utility is equal to the sum of component utilities?
2. Are component utilities stable over embedding conditions--two-, three-, and four-component menus? For example, is the utility for apple pie relatively the same in the context of roast beef alone as it is in the context of shrimp cocktail, roast beef and a menu price of \$6.50?

We first briefly describe the additive utility model and its relationship to this application. The results of the study are then presented, followed by a short discussion of the relevance of this type of research to managerial and consumer decision making.

Additive Conjoint Measurement

In additive conjoint measurement the researcher is interested in measuring jointly--at the level (asymptotically) of interval scales with common unit--an originally rank-ordered dependent variable and a set of independent variables. In so doing one assumes that a particular composition rule (viz., an additive one) describes the observed rank order. To illustrate, suppose a respondent is asked to rank a set of menus according to preference, each menu consisting of an appetizer, entree and dessert. Assume, further, that three different appetizers, entrees and desserts are available for making up the menus. Given three "levels" of each menu component, 27 distinct menu combinations are possible.

Each menu can be represented as an ordered \underline{n} -tuple:

$$(1) \quad x = (x_1, x_2, x_3)$$

where x_j ($j = 1, 2, 3$) denotes some value of the j -th factor. Other menu combinations are denoted x' , x'' , etc.

In additive conjoint measurement we seek real-valued functions on each factor (menu component) such that utility, $U(x)$, is expressed as:

$$(2) \quad U(x) = f_1(x_1) + f_2(x_2) + f_3(x_3)$$

and $x \geq_0 x'$ if and only if $U(x) \geq U(x')$, where \geq_0 is defined as an observable relation: "not less preferred than."

Thus, experimentally, one constructs factorial combinations of menu components, 27 combinations in the above case, and has the respondent rank order the combinations from most to least preferred.³ One then seeks a non-decreasing monotonic function of the observed ordering and part-worth functions, f_1 , f_2 and f_3 , whose values are added to produce $U(x)$, $U(x')$, etc.

Specifically, we can portray equation (2) in more traditional analysis of variance terms as

$$(3) \quad U(x) = \beta_0 + \beta_1 x_{21} + \beta_2 x_{31} + \beta_3 x_{22} + \beta_4 x_{32} + \beta_5 x_{23} + \beta_6 x_{33}$$

where β_0 denotes the contribution to utility when all factors are at their first levels and $\beta_1, \beta_2, \dots, \beta_6$ denote the incremental utility contributions of each factor at their second and third levels.

Additive utility theory has a long history (Becker & McClintock, 1967) but only recently have procedures been devised to measure component utilities in cases where the dependent variable is only rank ordered. One such algorithm, Kruskal's MONANOVA program (Kruskal, 1965), is used in this study.

Data Collection

There were 46 subjects who participated in this study: 23 males and 23 females, drawn from the Philadelphia area. The subjects were young adults in their mid-twenties to mid-thirties. The sample was drawn on a convenience basis. All interviewing was personally administered by trained student interviewers of the Wharton School.

Stimuli, Procedure and Tasks

The stimuli used in the study consisted of the names of 15 common food items--5 appetizers, 5 entrees and 5 desserts, as shown in Table 1. Also shown are the price levels (for restaurant-type menus) used in the last experimental task.

In the first phase of the interview the respondent was shown a set of twenty-five 3" x 5" cards. On each card was printed a pair of menu items--a specific appetizer and entree (denoted A-E in subsequent analysis). The subject was asked to imagine that he had been invited out to dinner by a friend and could order any of the 25 "menus" he desired (without being responsible for paying the bill). He was first asked to sort the cards into three ordered categories of overall liking and then to rank the cards within each category according to overall liking. (Cards could be shifted from category to category prior to final ranking.) This two-stage procedure was then repeated for 25 appetizer-dessert pairs (A-D) and 25 entree-dessert

Table 1

Stimuli and Background Variables Used in Menu Evaluation Tasks

<u>Appetizers</u>	<u>Entrees</u>
1. Shrimp Cocktail	1. Filet Mignon
2. Tomato Juice	2. Lobster Tails
3. Onion Soup	3. Roast Chicken
4. Fresh Fruit Cup	4. Roast Beef
5. Crabmeat Cocktail	5. Roast Pork

<u>Desserts</u>	<u>Menu Prices</u>
1. Jello	1. \$3.50
2. Apple Pie	2. \$4.25
3. Chocolate Pudding	3. \$5.00
4. Vanilla Ice Cream	4. \$5.75
5. Assorted Cheeses	5. \$6.50

pairs (E-D), each consisting of all appropriate combinations of the items shown in Table 1.

In the next phase of the interview the respondent was again shown a set of 25 cards. However, this time each card consisted of a triple of items--appetizer, entree and dessert (A-E-D). The combinations used for presentation were developed from a latin square design representing a one-fifth replicate of the full set of 125 factorial combinations. Instructions proceeded as before.

In the next phase of the interview the respondent was again shown a set of 25 cards. Each card contained a four component profile--appetizer, entree, dessert and price (A-E-D-P). The respondent was asked to imagine that this time he was responsible for purchasing his own meal. The combinations used for presentation were based on a graeco-latin square design representing a one-twenty-fifth replicate of the full set of 625 factorial combinations. Instructions for ranking proceeded as before.

Results

Each subject's ranking data regarding A-E, A-D, E-D, A-E-D and A-E-D-P were analyzed separately by means of Kruskal's MONANOVA algorithm. This program finds a best fitting monotone transformation of the original ranking and component utility scales such that the sum of the part worths of each menu component equals values of the monotonely transformed ranks. The program computes stress as a badness of fit measure. Stress represents (the square root of) a normalized residual sum of squares between additive model and monotonely transformed data values. A perfect fit of model to transformed

data is associated with a stress of zero. Stress values less than or equal to 0.1 are considered "good" fits from a rough, descriptive viewpoint.

In the case of tasks A-D, A-E and E-D a full factorial design was used. In tasks A-E-D and A-E-D-P the input data were treated as a fractional factorial and the missing data feature of MONANOVA was utilized. (Use of the latin square and graeco-latin square designs presuppose additivity of utilities.)

Additivity

Table 2 shows a frequency tabulation of stress by type of task.

Table 2

Frequency Tabulation of Stress Values from MONANOVA, By Task

<u>Stress Interval</u>	<u>A-E</u>	<u>A-D</u>	<u>E-D</u>	<u>A-E-D</u>	<u>A-E-D-P*</u>
$0 \leq \underline{S} \leq 0.10$	31	26	39	29	34
$0.10 < \underline{S} \leq 0.20$	9	8	2	7	8
$0.20 < \underline{S} \leq 0.30$	2	5	2	6	3
$0.30 < \underline{S}$	4	7	3	4	1
	—	—	—	—	—
	46	46	46	46	46
Mean <u>S</u>	0.077	0.120	0.049	0.102	0.064

* A -- Appetizer; E -- Entree; D -- Dessert; P -- Price

As can be noted from Table 2, the number of respondents displaying stress values of 0.1 or less for A-D, A-E, E-D, A-E-D and A-E-D-P, respectively, are 31, 26, 39, 29, and 34; that is, on the average, almost 70 per cent of the subject-tasks showed correspondence with the additive utility model.

Of course, high stress (poor fit) cases are equivocal. One cannot tell whether high stress reflects noisy data or a more complex, e.g., interactive-type, model. In the cases of the two component tasks, A-E, A-D and E-D, a full factorial design is represented. If, for the moment, one assumes that the original integer ranks can be treated as numerical (cardinal) values, standard ANOVA procedures can be used to examine the menu component interaction term. Accordingly, the original integer rank numbers for the A-E, A-D and E-D tasks

were submitted to separate ANOVA computations. (Since ranking data are used, each subject's mean and variance is constant; hence, sums of squares from this source of variation are zero.) Table 3 shows the results of the ANOVA calculations.

Table 3

Analysis of Variance Summaries for Two-Component Tasks

Appetizer-Entree (A-E)

<u>Source</u>	<u>Sums of Squares</u>	<u>d. f.</u>	<u>Mean Squares</u>	<u>F-Ratio</u>	<u>P</u>
Subjects -- S	0	45	0		
Appetizers -- A	1,051.14	4	262.78		
Entrees -- E	13,856.46	4	3,464.11		
S X A	17,449.66	180	96.94		
S X E	22,110.34	180	122.84		
A X E	126.18	16	7.89	1.09	> 0.05
Residual	5,206.22	720	7.23		
Total	59,800.00	1,149			

Appetizer-Dessert (A-D)

<u>Source</u>	<u>Sums of Squares</u>	<u>d. f.</u>	<u>Mean Squares</u>	<u>F-Ratio</u>	<u>P</u>
Subjects -- S	0	45	0		
Appetizers -- A	3,144.42	4	786.10		
Desserts -- D	4,683.49	4	1,170.87		
S X A	28,105.59	180	156.14		
S X D	17,212.90	180	95.63		
A X D	217.80	16	13.61	1.52	> 0.05
Residual	6,435.80	720	8.94		
Total	59,800.00	1,149			

Entree-Dessert (E-D)

<u>Source</u>	<u>Sums of Squares</u>	<u>d. f.</u>	<u>Mean Squares</u>	<u>F-Ratio</u>	<u>P</u>
Subjects -- S	0	45	0		
Entrées -- E	14,313.08	4	3,578.27		
Desserts -- D	4,335.09	4	1,083.77		
S X E	23,676.92	180	131.54		
S X D	13,022.91	180	72.35		
E X D	168.79	16	10.55	1.77	> 0.05
Residual	4,283.21	720	5.95		
Total	59,800.00	1,149			

Examination of the three ANOVA summaries of Table 3 shows that the contribution of sums of squares for menu component interaction to total sums of squares is indeed small. Although not shown in the table, the ratios of sums of squares due to menu component interaction to total sums of squares are only 0.002, 0.004 and 0.003, respectively, for A-E, A-D and E-D. Moreover, the largest mean squares are associated with either main effects (menu item-values) or subject-item interaction. None of the menu-item interaction terms is "significant" at the 0.05 level; however, in view of the ranking-type input data these "tests" are considered only descriptively.

Thus, given the reasonably low stress values and the small contribution of menu component interaction to total sums of squares (treating the integer ranking data as cardinal values), it would appear that the two-component menu data, at least, are rather well represented by an additive utility model. Of course, the possibility exists that some respondents, in some tasks, could be employing more complex (e.g., interactive) models.

Scale Stability Across Tasks

While it may be true that additive utility models provide good accounts of most subject-tasks, it does not necessarily follow that particular component utilities are the same across tasks. Accordingly, this question was considered next. For any specific subject the experimental design used here permits one to develop four interval scales each, for appetizer, entree and dessert items: a) the two scales for each menu component, as derived from the two-component tasks A-E, A-D and E-D; b) the scale derived from the three-component task; and c) the scale derived from the four-component task. Separate analyses can be conducted for appetizers, entrees and desserts.

To illustrate, for each of the 46 subjects, four scales were available for each of the 5 appetizers. At the individual subject level these scales can be intercorrelated in order to see if scale values differ across tasks. This was done for each subject in turn. Then an average correlation (of the six distinct two-variable correlations) was computed as a descriptive summary measure of how stable each subject's scales were across tasks.

The upper panel of Table 4 shows a frequency tabulation of the results of this analysis. We note that scale stability is greatest across tasks involving entrees, followed by appetizers and desserts.⁴ In all cases, however, the median correlation across subjects is 0.75 or higher.

The lower panel of Table 4, summarizing inter-task correlations (averaged over subjects), indicates that these correlations are also fairly stable.⁵ Again we note that stability across scales appears greatest in the case of entrees, followed by appetizers and then desserts.

Group Summary

In order to provide the reader with some idea of the substantive results of the study, MONANOVA scalings were also made of the ranks summed across the whole sample for each of the tasks separately. The part-worth scales found by MONANOVA were then translated and scaled, for comparison purposes, to unit sums of squares.

Figure 1 shows the results of these group analyses.

Table 4

Frequency Tabulation of Inter-Task Scale Correlations
By Subject and Task Pair

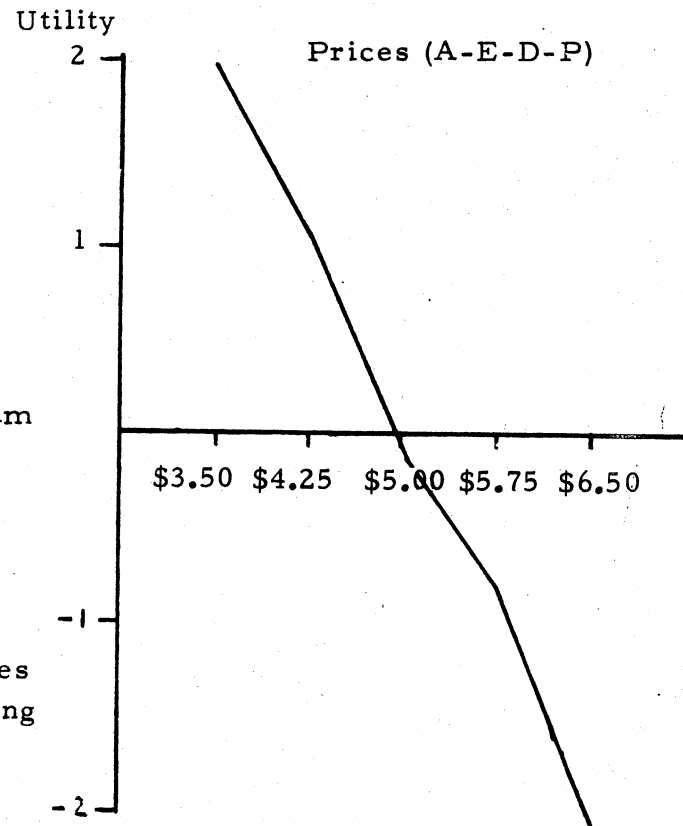
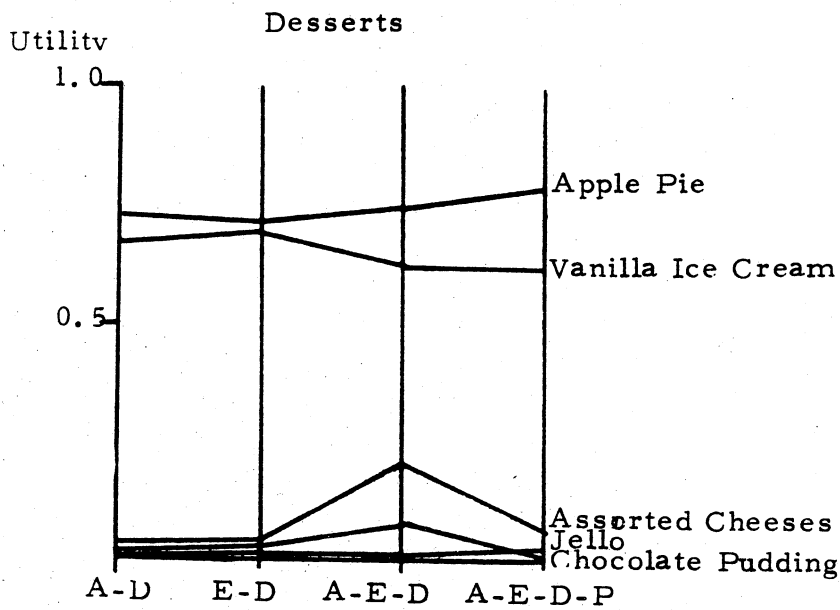
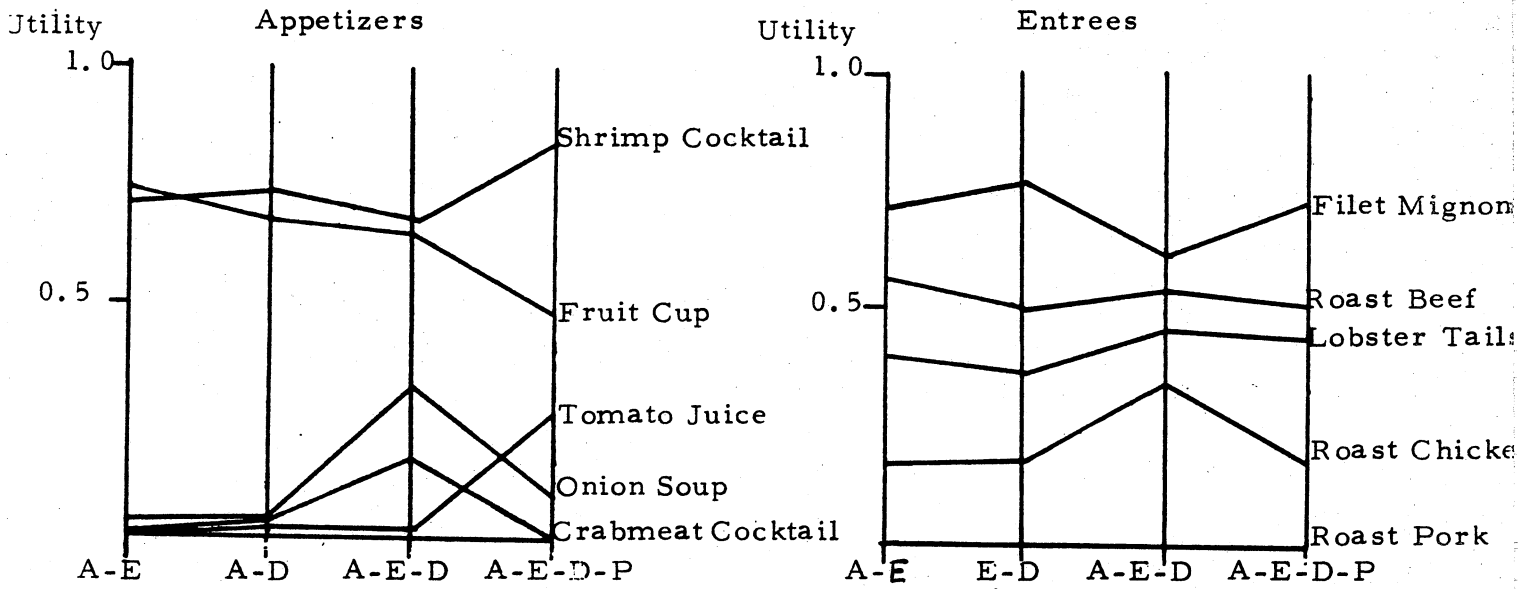
<u>Correlation Interval</u>	<u>Appetizers</u>	<u>Entrees</u>	<u>Desserts</u>
$0.8 < \bar{r} \leq 1.0$	23	37	21
$0.6 < \bar{r} \leq 0.8$	14	3	5
$0.4 < \bar{r} \leq 0.6$	1	2	14
$\bar{r} \leq 0.4$	8	4	6
	—	—	—
	46	46	46

Inter-Task Correlations (Averaged Across Subjects)

<u>Appetizers</u>		<u>Entrees</u>		<u>Desserts</u>							
<u>A-D</u>	<u>A-E-D</u>	<u>A-E-D</u>	<u>A-E-D-P</u>	<u>E-D</u>	<u>A-E-D-P</u>						
A-E	0.76	0.68	0.72	A-E	0.86	0.84	0.81	A-D	0.76	0.73	0.58
A-D		0.73	0.66	E-D		0.78	0.78	E-D		0.62	0.47
A-E-D			0.68	A-E-D		0.86	0.86	A-E-D		A-E-D	0.58

Figure 1

Part-Worth Functions (Normalized) by Component and Tasks



Scale stability at the group level differs somewhat by menu component. In the case of entrees all scale values are ordinally stable across tasks. We note that filet mignon is most preferred in all cases and roast pork is least preferred. Less stability is noted in appetizers, although the two highly favored appetizers, shrimp cocktail and fruit cup, remain well separated from the rest. Similarly, insofar as desserts are concerned, scale values lack stability although the two favored desserts, apple pie and vanilla ice cream, remain well separated from the remaining desserts across all tasks.

Finally, the behavior of price utility (as measured in the A-E-D-P task) in Figure 1 is of interest. As might be expected, component utility decreases with increases in menu price. Over the experimental range of this variable the relationship is almost linear.

Discussion

The foregoing report illustrates how conjoint measurement--more specifically, additive conjoint measurement--can be used to develop part-worth contributions of various components of a multi-component alternative to total utility. Moreover, these component utilities are (asymptotically) interval scaled with common unit, although the original dependent variable (the respondent's preferences for menus) is only rank ordered.

More elaborate utility models involving powers or interaction terms are capable of being applied within the conjoint measurement framework (Carroll, 1972). Indeed, conjoint measurement models exist for coping with responses that are expressed only categorically (Carroll, 1970), ordered or unordered.

The advantage of conjoint measurement procedures over other measurement methods lies primarily in their ability to deal with rank ordered response data. Moreover, as discussed by Krantz and Tversky (1971), conjoint measurement may be used to test alternative composition rules rather than seek scale values that are in best agreement with a prespecified composition rule (as was emphasized here).⁶ One might next inquire about the relevance of conjoint measurement to consumer and managerial choice behavior.⁷

Consumer Choice Behavior

A key consideration in the study of consumer choice is that most alternatives are multi-attribute. For example, current attitude theories postulate that total affect is functionally related to a set of "evaluative beliefs." Conjoint measurement provides a way to deal directly with the question of part-worth contributions to total utility (Becker & McClintock, 1967). That is, consumers can respond to the alternatives as total entities; it is the analyst who decomposes these responses according to some utility model, additive or otherwise (Einhorn, 1970). If the composition rule holds, the analyst should be able to predict choice among new combinations placed in the existing array of multi-attribute alternatives.

A number of measurement problems--price elasticity, package design, product benefits analysis, perceived risk--would seem to be amenable to conjoint measurement designs although little empirical research has been conducted on these problems so far.

Managerial Choice Behavior

The study of managerial decision making is also characterized by multiple objectives--profits, growth, market share, environmental effects, and so on. Somehow managers "balance" conflicting objectives by means that are rarely explicated to the researcher. Here, too, it would seem that various utility models might be explored via conjoint measurement, including the study of self-explicated models (Hoepfl & Huber, 1970). With growing awareness of the complexity of modern decision making, it seems clear that future research attention will be increasingly paid to multi-attribute choice problems in both the private and public sectors. Conjoint measurement could play a methodological role in the measurement tasks associated with these problems.

Footnotes

1. Paul E. Green is S. S. Kresge professor of marketing, Yoram Wind is associate professor of marketing and Arun K. Jain is a doctoral student in marketing, all in the Wharton School, University of Pennsylvania. Financial support, in part, was provided by the General Electric Educational Foundation.
2. The procedure and analysis described here represents only part of the complete study, a full account of which can be obtained from the authors.
3. If the number of factor-level combinations is too large for convenient ranking one can use fractional factorial designs of the type employed in this study.
4. The relative stability of scales was in accordance with the relative importance of each component class, as independently reported by each respondent. That is, each respondent was also asked to rate (on an 11-point scale) the relative importance of each menu component to overall meal satisfaction.
5. However, some tendency is noted for correlations to decrease as the embedding context becomes more complex, e.g., three-component or four-component versus two-component menus.
6. Unfortunately, appropriate statistical tests for dealing with fallible response data (where composition rules are not perfectly satisfied ordinally) are still lacking.
7. See Green and Rao (1971) for a discussion of some of the current limitations of conjoint measurement.

References

- Becker, G. M. & McClintock, C. G. Value: Behavioral Decision Theory. Annual Review of Psychology (1967) 239-86.
- Carroll, J. D. Categorical Conjoint Measurement. Multilithed report. New Jersey: Bell Telephone Laboratories, 1970.
- Carroll, J. D. Individual Differences in Multidimensional Scaling. In Shepard, R. N., Romney, A. K. and Nerlove, S. (eds.), Multidimensional Scaling: Theory and Applications in the Social Sciences. New York: Academic Press, 1972.
- Einhorn, H. J. The Use of Nonlinear, Noncompensatory Models in Decision Making. Psychological Bulletin 1970, 73, 221-230.
- Green, P. E. & Rao, V. R. Conjoint Measurement for Quantifying Judgmental Data. Journal of Marketing Research, 1971, 8, 355-63.
- Hoepfl, R. T. & Huber, G. P. A. Study of Self-Explicated Utility Models, Behavioral Science, 1970, 15, 408-14.
- Krantz, D. H. & Tversky, Amos. Conjoint Measurement Analysis of Composition Rules in Psychology, Psychological Review, 1971, 78, 151-69.
- Kruskal, J. B. Analysis of Factorial Experiments by Estimating Monotone Transformations of the Data. Journal of the Royal Statistical Society, Series B, 1965, 22, 251-263.

CONSUMER JUDGMENT STRATEGIES:
BEYOND THE COMPENSATORY ASSUMPTION

Peter L. Wright¹
University of Illinois

An important problem which is receiving increasing attention among consumer researchers focuses on consumer judgment of competing alternatives. While historically there has been interest in the processes of judgment and choice in the marketplace, the acceleration of social and product innovations which will characterize tomorrow's marketplace (which has provoked one contemporary social commentator (Toffler, 1970) to predict a society characterized by "over-choice") has heightened the need for rigorous examination of different models of consumer judgment strategies. Recent consumer research attention has centered on a linear additive "attitude" model as providing insight into consumer choice behavior. The models proposed by Fishbein (1964) and Rosenberg (1956) have been most popular; these models belong to a class of models which may collectively be labeled "compensatory-processing" models. These models have in common the identification of two variables as important in mediating a consumer's behavior toward a product; beliefs about the product characteristics and the evaluative importance of the characteristics. The basic assumption is that individual consumers affectively evaluate alternatives along a set of dimensions. The consumer is then seen as combining these unidimensional cues according to a simple linear rule. The "index" resulting from such a process is expected to guide the individual's choice according to a decision rule which says "choose the alternative with the highest index." Attention has been directed at demonstrating that this two-component linear model satisfactorily predicts behavior (Bass and Talarzyk, 1972; Hansen, 1972), and in operationally refining measurement of the two variables (Cohen and Houston, 1970; Schendal, Wilkie, and McCann, 1971).

There has been little interest thus far in critical examination of the assumption about the strategy an individual consumer may use in combining information about a multidimensional product which is implicit in the additive compensatory model (however, see Bettman, 1971, and Russ, 1971). One reason for this undoubtedly is that this assumption has been the dominant one across disciplines where human choice behavior has been a topic of empirical inquiry. The general success of predictions based on such linear, compensatory models cannot be disputed, and the search for ways to detect the actual operation of some choice strategy other than the linear one among a variety of choice alternatives has not been overwhelmingly successful.

Two comments should be made however in order to place the continued pursuit of non-compensatory choice strategies in perspective. First, it is generally agreed that the formal linear model is statistically extremely powerful, and is capable of providing close approximations to data which is not itself strictly linear (Green, 1968). We must keep in mind the important distinction, however, between successful prediction of consumer preferences and successful representation of the actual cognitive processes which individual consumers are bringing to bear on the choice task. Prediction is relevant to certain of the objectives of the publics we serve (managers and consumer advocates), but prediction alone frequently is accompanied by only a low level of understanding of the psychological behaviors involved. Second.

it will probably prove more useful in the long run to begin to directly attack the question of when an individual might use different available strategies rather than trying to prove the general superiority of one or another model.

Alternative Consumer Judgment Strategies

The linear compensatory model has been verbally described above; its formal presentation takes the form:

$$U = \sum w_i s_i$$

where: U is the overall rating or ranking of the alternative
 w is the weight assigned to characteristic i
 s is the evaluative (affective) rating of characteristic i

Other models of the judgmental process have been proposed by social-psychologists (e.g., Coombs, 1964; Dawes, 1964). One such model, labeled the "conjunctive" model, suggests that an individual uses a minimum evaluation procedure in which the multidimensional object is judged on its minimum performance on all characteristics. The contrast between this strategy for judging the product and that assumed by compensatory strategy should be understood. When using the compensatory strategy, the consumer focuses equal attention across the entire range of possible ratings for product characteristics, and allows a better than average performance on one characteristic to effectively cancel out a poorer than average performance on another characteristic. If the individual is using a conjunctive strategy, he focuses a disproportionate amount of attention on the negative end of the evaluative continuum. Assume the individual cognitively arrays product characteristics on seven-point bipolar scales where 1 = the worst rating possible and 7 = the best rating possible. In terms of this scale, when using the conjunctive strategy he tends to ignore ratings falling at the number 5, 6, or 7 points relative to those falling at 1, 2, or 3.

An alternate strategy, the "disjunctive" model, suggests that the consumer may judge the product on its best characteristic regardless of the other attributes. This strategy is essentially the converse of the conjunctive strategy, since now a disproportionate amount of attention is focused on the upper (positive) end of the continuum for evaluation. Points 7, 6, and 5 are more influential in the overall judgment than the remainder of the scale.

Until recently suitable functions to represent mathematically the conjunctive and disjunctive strategies were not available. Einhorn (1970) has noticed that two fairly familiar geometric surfaces can serve as reasonable approximations of these models. He argues that the conjunctive model can be represented by the parabolic response surface:

$$U = \prod_{i=1}^n (s_i)^{w_i}$$

or

$$\log U = \sum_{i=1}^n w_i \log (s_i)$$

Here the consumer's judgment of the multidimensional product will be highest

where there are no wide discrepancies in ratings across the separate product dimensions.

The disjunctive strategy is approximated by a hyperbolic response surface, and the following mathematical equation is appropriate:

$$U = \pi \prod_{i=1}^n [1 / (a_i - s_i)]^{w_i}$$

or

$$\log U = \sum_{i=1}^n -w_i \log (a_i - s_i)$$

where: U = overall product rating or utility

s_i = evaluative ratings of characteristic i

w_i = weight of characteristic i

a_i = some arbitrary value set above the highest s_i rating (i.e. above the asymptote) so that U does not become infinite

Here the consumer will find the product quite attractive if he can rate it extremely high on a single attribute (of course if he can rate it extremely high on more than one attribute, he is ecstatic).

Einhorn points out that although the linear model appears to be mathematically simpler, its actual use by the individual may in fact require considerably more cognitive exertion than either of these two alternate strategies which require more complex mathematical equations. There is considerable precedence, it should be noted, throughout the study of behavior for researchers to adopt as their working model of their subjects' behavior one which is easy to analyze from the perspective of the researcher, regardless of whether the model would in practice be simple and efficient for the subjects to use.

Individual Use of the Strategies

One purpose of this paper is to present data demonstrating that individuals do indeed appear to adopt strategies in judging alternative products which can be better represented by the non-linear, non-compensatory models than by the traditional linear model. The method of inquiry was to require individuals, in this case males, ages 19-22, who did not at the time personally own an automobile and were approaching a change in financial status which made them likely purchasers during the upcoming year, to rank order two sets of fifteen hypothetical automobiles. These cars were described for the subject in terms of their rating on four attributes: selling price, style, ease of handling, riding comfort. The information about each attribute was given on a seven-point scale which varied from "greatly above average" to "greatly below average," with the middle point "average." All the cars in both stimulus sets were described as belonging to the "under \$4000" class, and the ratings referred to this particular class of cars.

Subjects were presented with the first set of hypothetical cars with each one described on a separate card. They were asked to rank them in order of personal preference as products which they themselves would purchase upon graduation. Having ranked the first set, they repeated the judgments for a second validation set.

For each individual subject, a linear regression analysis was used to determine the weighting parameters. The use of regression coefficients as approximations of the weighting of the product dimensions in the judgment exercise is, it might be noted, more clearly justified when, as was the case here, the individual is the unit of analysis than when the weights derive from the more common aggregated analysis. In the case of the conjunctive and disjunctive models, the weights were found by using multiple regression on the log-transformed variables. The weights obtained from the first set of automobile stimuli were then applied to the second set in order to obtain a predicted ranking. The actual ranking of the second set by each subject was then correlated with the derived predicted ranking by means of the Spearman rho procedure. These second set rankings were used as the criterion for comparisons with the predicted rankings from the linear, conjunctive, and disjunctive models.

The data for two of the subjects will be discussed. For Subject A the rank ordering predicted by the linear model correlated .68 with the actual ranking; the ranking predicted by the conjunctive model correlated .89 with the actual ranking; and that predicted by the disjunctive model correlated .26 with the actual ranking. The difference between the linear and conjunctive models is significant at the .01 level, as is the difference between the linear and disjunctive models. It appears that this subject's judgment strategy was best represented by the parabolic function (conjunctive model), and that the disjunctive model was distinctly inappropriate.

Subject B appeared to approach the judgment task differently than Subject A. In his case, the ranking predicted by the linear model correlated .55 with the actual ranking; the ranking predicted by the conjunctive model correlated .57; and the ranking predicted by the disjunctive model correlated .82. The difference between the correlation of the disjunctive model and that of either of the other models is significant at the .05 level. The strategy of Subject B was apparently more closely approximated by the hyperbolic function associated with the disjunctive model than by the linear or parabolic functions.

For both these subjects, examples can be found in the data to illustrate where either the conjunctive or disjunctive models picked up fairly dramatic errors in the predicted rankings obtained using the linear model. These data have been selected in order to clearly demonstrate that individual consumers may use a variety of rules for psychologically combining separate pieces of information to arrive at an overall product judgment.

The Need for Systematic Study

There appear to be compelling reasons for expanding the study of consumer judgment and choice to include the probability that individual consumers employ various strategies other than the monolithic linear compensatory strategy: (1) it can be successfully demonstrated, as the data presented show, that psychometric functions which are theoretically justifiable representations

of the conjunctive and disjunctive strategies provide a better fit to the observed judgments of some individual consumers than does the linear model; (2) the subjects themselves, when asked to verbally recapture the nature of their own just-completed judgment process, will frequently describe a conjunctive or disjunctive strategy quite clearly (Goldberg, 1968). This was observed in the present study even in subjects for whom the formal conjunctive or disjunctive models afforded no predictive advantage relative to the linear model; and (3) the importance of the point raised earlier about the amount of cognitive exertion required of an individual consumer in applying these strategies to process product information cannot be minimized. Much of what we know about cognitive behavior leads us to expect that people will use devices for simplification of their information handling tasks wherever possible; this is an efficient adaptation to their complex environment which enables them to buy time for a host of other activities. Perceptual categorization is one example of this. Realistically, for a consumer to actually do what the linear, compensatory model says he does would be much more difficult and take much more time than doing what the alternate models, with their "focusing-on-one-end-of-the-evaluative-dimension" property, suggest.

Relevance to Research on Dynamics

If in fact individual consumers do use various judgment strategies, then another important reason for systematically studying this area springs from what we are ultimately about as consumer researchers. Much marketing activity is concerned with influencing, rather than merely predicting, consumer behavioral preferences. Thus while researchers may start with structural models of established predictive validity, the ultimate value of such models appears to be in bringing order to the study of how and why preferences are changed by changes in product features, advertising, distribution, etc. Currently the study of attitude change is beset by problems of inaccurate predictions and difficult to interpret findings. Consider the relevance of the question of alternative judgment strategies. Predictions as to what happens to a person's overall judgment of a product when a belief about some single characteristic is changed are entirely dependent on which judgment strategy the person is using relative to that product. An example may clarify this idea. Assume Consumer X has formed a judgment of a product based on his beliefs about three characteristics of the product. We now wish to predict the effect on his summary judgment of a slight change in his rating of the product on one of the three dimensions. (Assume he thinks each dimension to be equally important; this will simplify the discussion although it is not necessary to the argument.) If Consumer X is indeed using a linear, compensatory judgment strategy in this case, then a change of one scale-unit on any of the three belief dimensions will have an equivalent effect on the overall judgment. This is true regardless of what his prior beliefs were, i.e., where he previously located the product on each of the belief dimensions. However, if Consumer X is instead applying a conjunctive strategy, our expectations about the outcome of such changes becomes contingent on both the prior belief and the specific dimension where change is accomplished. Since the conjunctive strategy implies that the judgment will not improve until the product reaches satisfactory cutoffs on each dimension, a positive change in an already satisfactory belief will not make any difference. The overall judgment will show a shift only when the change is accomplished along the dimension where prior beliefs fell below the minimum cutoff. Similarly, if Consumer X is applying a disjunctive judgment rule here, then changing previously negative beliefs to neutral or neutral beliefs to mildly positive

will not have a noticeable effect on the overall judgment. Only by creating the belief that the product is superior on one (or more) dimensions can the judgment improve.

For the linear model, all locations on the belief dimension are equivalent. For the conjunctive or disjunctive models, certain locations are markedly more important than others to the consumer. These differences become extremely important in setting objectives for an advertising program, in stating researchable hypotheses, and in trying to explain why a certain manipulation was successful at the belief level but not at the attitudinal level. Considerations of the possibility of alternate judgment rules can no longer be ignored in research and theory on attitude dynamics.

Modeling the Environment for Consumer Judgment

As has been seen, individuals differ in their use of strategies in the same task. There are also undoubtedly differences within the same individual across various tasks. One further purpose of this paper is to outline a conceptual model of the task environment, specifying certain dimensions of the situation in which the judgment occurs which may influence which model a consumer applies.

Perception of Penalties and Rewards

Consumers may be aware that the situation they are facing places particularly heavy penalties on making either a Type I or Type II error in their judgment of the product. If such is the case, they will probably adapt their judgment conventions to accommodate this analysis much as researchers have done in their setting of conservative significance levels to protect against certain errors in judging empirical evidence. If the consumer perceives that the probability that he will be drastically penalized for committing an error of inclusion is high, he should favor the conjunctive strategy. If the housewife expects to be blamed for her mistake in choosing a product (and not rewarded significantly for especially good choices), the conjunctive strategy is quite reasonable. The same applies to the industrial buyer's perception of reward system in his organizational role. Conversely, if the individual believes that he will be penalized for excluding a product which is extraordinarily satisfying, he may favor a disjunctive strategy. Such penalties will likely be in the form of opportunity costs: praise or tangible rewards forgone.

The individual may perceive these penalties to apply only for the specific product, for all product decisions made in a particular role, or for most decisions he has and will encounter in life. When a person tends to develop a chronic strategy for making judgments based on this latter type of perception, then his judgment style takes on the properties of a personality trait. It would not be surprising if the general environment itself is set up to favor certain strategies; for example it may be that penalties for errors of inclusion are simply much more prevalent in society than penalties for exclusion. Or, if not more prevalent, than perhaps more uncomfortable psychologically, especially in the short run.

Exclusiveness of the Behavioral Expression of the Judgment

Presumably a person's judgment of a product will carry with it certain behavioral ramifications. The futility of measuring predispositions which

are not linked to specified behaviors has been noted. Sometimes the context of the judgment will be such that the person must choose one and only one of the alternative products available, and the selection is not revocable. In other cases, the behavioral choice may not be mutually exclusive of the choice of other alternatives. For example, a person may have a given set of needs to satisfy with a fixed resource. If he can only acquire a single product alternative, given this limit, he will tend to favor a conjunctive or linear strategy since these arrive at their final judgment through some consideration of performance on all dimensions. If he has, however, the luxury of being able to select more than one of the alternatives, he may use the disjunctive strategy. Here he could look for outstanding performance on one attribute of Product A to satisfy a certain need quite well, outstanding performance on another attribute of Product B to satisfy another need quite well, and so forth. Again, the question of penalties is relevant, and this factor can be considered a subset of the above discussion.

Range of Alternatives Being Judged

It is quite likely that the strategies may be used sequentially. The aim of all product judgments is eventual selection of some appropriate behavior toward the products. At a certain stage in selecting appropriate behaviors to satisfy certain needs, the consumer will find that there are a large number of candidate products which he must judge. It would be efficient at this stage to apply the conjunctive strategy to whittle down the domain of alternatives (hopefully) to a more manageable group. It was noted earlier that application of the linear model is quite complex, and its application to each product in a large group is even more exhausting. Having eliminated many of the candidates as worthy of positive behavioral action, the consumer may then be able to apply the linear model as the discriminator. Or he may first use the conjunctive to eliminate all products not meeting the cutoff, then the disjunctive to eliminate all not having at least one superior characteristic, then the linear. This corresponds to an attempt at cognitive simplification.

Of course, the use of differential weighting of the attributes involved is another such device for simplification. In the case above, one weights different parts of the dimensions differently to make judgments easier, but assigning greater importance to specific dimensions as a whole is also a commonly recognized method for accomplishing the same thing. (This is another case where the subject complicates the researcher's formal model in order to make things less complicated for himself.) Thus, by weighting one or two dimensions very heavily and others so lightly as to virtually ignore them, a choice may become apparent. If the individual is not able, however, to really reduce the set of evaluative product dimensions he feels worth considering to one or two, then this type of simplification doesn't work for him.

Does It Work?

An ultimate question which the consumer confronts in situationally applying one or the other judgment strategies is whether or not that strategy does indicate for him the product alternative which he should choose. After all, that is his goal in going through the process of integrating all this information. If, in judging a number of alternatives, he can not discriminate among them using one strategy he will undoubtedly try using another until at last some choice is indicated. If this is impossible (and there are certainly

product domains where there exists sufficient homogeneity among the alternatives such that all of these schemes are frustrated), then additional information must be gathered to enable a choice. In terms of the sequential pattern of use discussed above, if the consumer faced with a lot of alternatives to evaluate applies any of the models and, lo and behold, finds that one discriminates so successfully that no further evaluation is necessary, then the process stops right there.

Working Forward vs. Working Backward

One final distinction might be made which is perhaps more closely related to the conduct of research on this topic than to the behavior of the consumer. Two research paradigms used in the study of judgment and attitude models must be contrasted. One approach--that illustrated here--has the subject respond to hypothetical stimuli described for him by the researcher. In such a case, the subject presumably does not already have any attitude formed toward the stimulus prior to receiving the experimental information, and his beliefs about the stimulus characteristic are controlled by the researcher. His task is clearly to integrate the information and arrive at a judgment. The contrasting approach is to utilize as the stimulus some real and identifiable object with which the consumer individual has had previous experience. Here, the subject enters the research situation with his judgment already formed and the latent structure of beliefs is extracted from him by the researcher. In both cases, different models may be fit to the resulting data and their relative predictive validity examined.

These two paradigms are actually studying different phenomena. The former is clearly focusing on the development of a judgment from its component parts, whereas in the latter this creative aspect is missing. It may very well be that these differences in the actual behavior being studied do have a biasing effect on the probability that a certain model will appear superior. It would appear, for example, that the situation in which the subject works his way backward from his existing attitude to his separate beliefs will tend to promote a linear strategy. Just as when, in originally integrating the information, it may have been simpler to adopt a non-linear strategy, when describing the structure behind an already developed attitude it may be simpler for the subject to cognitively run a linear gradient through the belief-scales he is filling out.

The whole point in a person developing an attitude in the first place is that it frees him from constantly having to go through the process of re-integrating all the information he carries about the product whenever some potential for acting toward the product occurs. Thus the individual asked to describe his belief system may often be forced to recreate it in part. In any case it is clear that the two research situations do differ in terms of the direction of the dynamics involved, and that there is no necessary reason for the processes of someone working backward to mirror the processes involved in integrating new information.

Footnotes

¹Peter Wright is Assistant Professor of Business Administration at the University of Illinois, Urbana.

References

- Bass, F. W. and Talarzyk, W. W. An attitude model for the study of brand preference. Journal of Marketing Research, 1972, 9, 93-96.
- Cohen, J. B. and Ahtola Houston, M. The structure of consumer attitudes. Working paper, University of Illinois, 1971.
- Bettman, J. R. The structure of consumer choice processes. Journal of Marketing Research, 1971, 8, 465-471.
- Coombs, C. H. A Theory of Data. New York: Wiley, 1964.
- Dawes, R. M. Social selection based on multidimensional criteria. Journal of Abnormal and Social Psychology, 1964, 68, 104-109.
- Einhorn, H. J. The use of nonlinear, noncompensatory models in decision making. Psychological Bulletin, 1970, 75, 221-230.
- Fishbein, M. An investigation of the relationships between beliefs about an object and attitude toward that object. Human Relations, 1964, 16, 233-240.
- Green, B. Descriptions and explanations: a comment on papers by Hoffman and Edwards. In B. Kleinmütz (Ed.), Formal Representation of Human Judgment. New York: Wiley, 1968, 91-98.
- Goldberg, L. Simple models or simple processes? some research on clinical judgements. American Psychologist, 1968, 23, 483-496.
- Hansen, F. Consumer Choice Behavior: A Cognitive Theory. New York: The Free Press, 1972.
- Russ, F. Evaluation process models and the prediction of preference. Proceedings, Second Annual Conference, Association for Consumer Research, 1971, 456-461.
- Scandal, D., Wilkie, W., and McCann J. An experimental investigation of attribute importance. Proceedings, Second Annual Conference, Association for Consumer Research, 1971, 404-416.
- Toffler, A. Future Shock. New York: Random House, 1970.

EFFECTS OF THE NUMBER AND TYPE OF ATTRIBUTES INCLUDED IN AN
ATTITUDE MODEL: MORE IS NOT BETTER

William L. Wilkie and Rolf P. Weinreich
Purdue University

In recent years, there has been a growing interest in marketing applications of multi-attribute attitude models of the form originally proposed by Rosenberg and Fishbein [7,2]. In behavioral theory, an attitude is defined as the evaluative dimension of a concept - a relatively stable affective response to an object [3,6]. Specifically, in Fishbein's formulation, attitude formation is viewed as a process of cognitive summation. Essentially, the principle leads to the prediction that an individual's attitude toward any object is a function of his beliefs (i.e., perceptual or cognitive constructs) about the object, and the evaluative aspects of those beliefs:

$$(1) \quad A_o = \sum_{k=1}^s B_k \cdot a_k$$

where, for each individual,

A_o = the attitude toward object "o"

B_k = the strength of belief k about o

a_k = the evaluative aspect of B_k

s = the number of beliefs

As can be seen from the formula, an attitude is the weighted sum of beliefs about an object, the weights representing the relative contribution of each belief in forming an attitude. Typically, model performance is tested by correlating the derived attitude scores with an independent measure of affect.

In the marketing context, this approach has been extended to postulate that attitudes toward brands are governed by a consumer's beliefs about the ability of different brands to satisfy specific product attribute intensities he desires. Although the situation is somewhat different, in that several "attitude objects" are evaluated, a similar model structure has been proposed to operationally determine an individual's attitudes toward each of j brands.

$$(2) \quad A_j = \sum_{k=1}^s I_k \cdot B_{jk} \quad \text{model I}$$

where, for each individual,

A_j = the attitude toward brand j

B_{jk} = the rating of brand j on attribute k

I_k = the importance of attribute k in forming
an overall attitude toward brands

Instead of predicting affect for a single concept, the model is used to predict the preference orderings (or other external criteria) for competing brands. Thus, model performance is tested by correlating a set of attitude and preference ratings:

$$(3) \quad P_j = f(A_j) = f\left(\sum_k I_k \cdot B_{jk}\right)$$

where for each respondent,

P_j = the preference for brand j

A_j = the attitude toward brand j

Some thirty marketing studies related to this basic model have been reported in the past three years, reflecting the natural appeal of the multi-attribute notion to marketers. If attitudes are intervening, evaluative responses - indeed removed from, but closely associated with choice - then the model might provide operational measures of how a brand is being perceived and evaluated vis à vis its competitors. This information might then be useful for brand positioning decisions.

While considerable testing of the predictive and diagnostic power of this approach has been undertaken by marketing, several issues clouding its practical application remain unresolved [8]. One important issue area concerns the nature and handling of attributes themselves. Attributes ("k" in formula 2) are used as basic dimensions of the model; respondent ratings of the importance of each attribute and brand possessions of the attribute then provide the raw data for calculations of brand attitude scores for each individual. There are two broad decisions involved in the model's use of attributes:

- (1) initial specification
- (2) inclusion in the calculations of brand attitude

Initial specification concerns the manner in which a candidate list of product attributes - typically representing perceptual rather than concrete product characteristics - is generated by the research. Marketing applications have usually utilized some form of unstructured pretests to generate lists of 5 - 30 attributes which are then presented to respondents for brand belief and importance ratings.

"Inclusion in the model" refers to the manner in which the raw data is used by the researcher. In particular, are all attributes in the initial list applicable to all respondents? If not, how many and which attributes should be used for each respondent? Typical marketing studies have used all attributes for every individual and have thus assumed that the same belief structure holds for all consumers.

This paper reports the results of a study comparing the efficacy of this common assumption versus analyses which explicitly manipulate the number and type of attributes included so as to better reflect individual differences in attitude structures. A brief review of relevant literature on this question is followed by descriptions of the data, methodology, and results of this study. It is argued that neither theory nor prior research support the practice of including all attributes for every respondent; this paper demonstrates the magnitude of some drawbacks inherent in this practice and offers feasible procedures which allow for idiosyncratic attitude structures within the framework of a multi-attribute model.

Discussion of Past Research

There is little explicit argument of this matter to be found in the literature, but a careful reading of several prominent papers can shed some light on the issue.

One of the earliest accounts of the summative model was presented in social psychology by Rosenberg [7]. His model structure is similar to formula (1), although its theoretical underpinnings are somewhat different to those later postulated by Fishbein. Rosenberg tested his model by using the summative form on 35 value items related to "whether or not members of the Communist Party

should be allowed to address the public." He concluded that "beliefs associated with an (independent) attitudinal affect tend to be congruent with it, i.e., that there exists within the individual an 'organization' of the affective and cognitive properties of his total pattern of response to what, for him, is an 'attitude object.'"

This statement bears the clear implication that individual differences in the type and number of beliefs should be recognized. Although Rosenberg tests the effect of utilizing different numbers of values in the summation model, it is not clear in his report whether the composition of values was allowed to vary for each individual. In a second test using only idiosyncratic "salient" beliefs, however, Rosenberg notes that predictive power increased as compared to using all thirty-five values.

Fishbein defined salient beliefs to be "those present in the individual's response hierarchy" [2] and noted that inclusion of nonsalient beliefs will tend to reduce the precision of attitude measurement. Again, the implication is that individual differences in the number and type of "salient" attributes should be recognized. However, a later model test [1] uses a common number and types of beliefs for each respondent - a seeming contradiction with Fishbein's original postulates about individual differences in attitude structure.

Since the context for marketing differs from the concerns of social psychology, it is possible that the proposals of Rosenberg and Fishbein might not be of pragmatic value for the attribute issue. In marketing, there seems to be agreement that attributes are likely to be product-specific, that reasonable attributes will generally be small. But the marketing literature also lacks discussion about the issue of idiosyncratic attribute inclusion. Following Fishbein's position, the virtually unanimous stand taken is that only "salient" attributes should be included in the model. Most papers assume, however, that the original list of attributes (generally 5-7) is sufficiently and exhaustively "salient" for all respondents. Operational definitions of salience are rarely undertaken. Equating salience with "importance," Hansen [4] presents one of the few evidences of the effect of attribute types included in the model. He chose the three most important attributes for each respondent, compared model performance against utilizing all attributes, and noted that predictive power was not diminished by using fewer attributes. However, no test on the effects of varying attribute numbers was mentioned.

This paper moves beyond traditional testing of the model by systematically looking at the effect of attribute inclusion in the basic summative model. The hypothesis advanced is that gains can be achieved when individual differences are allowed in the number and type of dimensions utilized in computing a respondent's attitude toward brands. Results of using varying numbers of attributes from a common list, of allowing the types of attributes included to vary (holding their number constant), and of permitting both type and number to differ from respondent to respondent are compared against the traditional approach of utilizing all ratings for each individual.

Methodology

The Data

A convenience sample of twenty-nine housewives living in a small Eastern town were asked to rate seven supermarkets on seven store attributes (using seven-point bipolar scales): location, prices, quality, variety, closeness to other services, shopping climate, and "other shoppers." The attribute set was generated from a more extended list, used in preliminary interviews. In addition, subjects were also asked to rate the importance of the attributes in making a choice of a shopping place, as well as to rank preferences for the seven stores in the area.

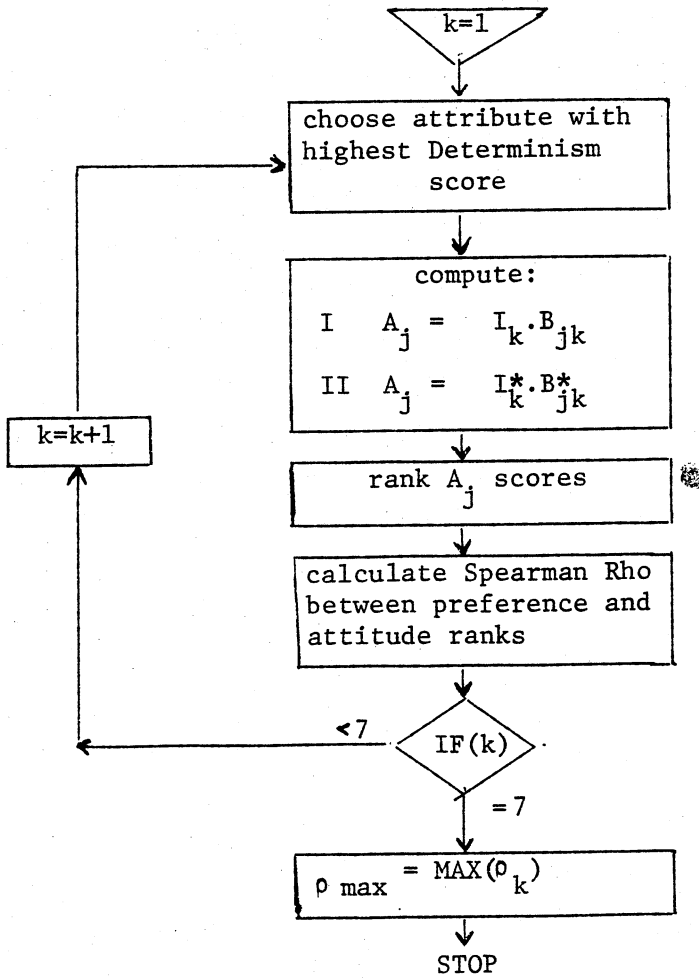
Analysis

The traditional approach to these data is represented by applying the seven store attributes to all subjects, computing an attitude score for each respondent using formula (2), ranking the derived scores, and correlating attitudes with preferences using Spearman's Rho index to determine the degree of association of the two measures.

The approach in this paper is methodologically different in that attributes are entered into the summation in stages. According to some criterion or entry rule (to be discussed shortly), one additional attribute is entered at a time, until all seven have been included. At each summation stage, the basic formula is applied to compute attitude scores for each brand and each individual. These are then correlated with stated rank preferences. Figure 1 shows in flow-chart form the procedure used.

Figure 1

Description of the Staged Entry Approach



Spearman Rho correlations between ranked attitudes and preferences are used to measure performance of the model. Thus, by comparing correlations at each stage with correlations represented by the traditional approach, changes in model performance can be analyzed with respect to changes in inferred attitude dimensionality.

Criteria for Attribute Entry

Given a list of attributes the traditional approach has introduced all ratings to the model; the order in which attributes enter is of no import. When staged inclusion is attempted, however, the order of entry is of obvious concern, and some rule or criterion is required to ensure that "salience" is represented in entry. As noted, little discussion has been devoted to the definition and operationalization of "salience" in previous marketing studies. Two allied constructs have, however, been advanced: importance and determinism. The study reported here utilized both rules as criteria with which to generate order of attribute entry to the model; results for the determinism criterion are reported.² Determinant attributes, as proposed by Myers and Alpert [5], are those "which are most closely related to preference or actual purchase decisions." Car safety is an example of a dimension which is probably important, but not determinant for most consumers due to perceptions that brands do not significantly differ on this attribute. Thus, the Determinism concept incorporates both importance weights and the notion of disparity in perceived satisfaction on the attributes through different brands. Operationally:

$$(4) \quad D_k = I_k^* \cdot \sigma_k$$

where, for each respondent,

D_k = the determinism score for attribute k

I_k^* = the standardized importance score for attribute k

σ_k = the standard deviation (over all stores) of attribute k

Four studies were performed for comparison purposes:

1. The first represents the traditional approach, whereby all attributes (7 here) are included for each individual in the summation formula, and correlations with stated preference are computed.
2. The second study measures the impact on model performance of allowing the number of attributes to vary from 1 to 7, but constraining them, at each stage, to be of the same type. The mean Determinism score for each attribute across the sample was used to order entry. Thus, attitude scores were computed for all individuals after the most determinant attribute; the two most determinant attributes (and so on) had been entered into the summation, until all seven had been included. By comparing correlations at each stage with the traditional approach, the degree to which a smaller set of attributes (<7) performs better, equal or worse can be analyzed.
3. A third study concentrated on the effect of attribute types on model performance. To derive attitude scores at each stage, attribute inclusion was resolved by using each individual's Determinism scores on each attribute as entry criteria. This analysis allows a study of the extent to which individual differences in the type of dimensions (but using the same number) will result in better performance than using a common set (i.e., study 2).

Table 1

Results from the Staging Approach
for Two Selected Respondents

Respondent # 11

Model form: $A_j = \sum I_k \cdot B_{jk}$

Order of Attribute
Entry

	1	3	2	5	4	6	7
Determinism score	2.83	.43	.36	.32	.02	-.33	-3.83

<u>stage</u>		<u>store number</u>							<u>RHD</u>
		<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	
1	attitude score	60	10	40	10	70	70	50	.61
	score ranks	3	6.5	5	6.5	1.5	1.5	4	
	prefer. ranks	3	5	2	7	1	4	6	
2	attitude score	92	34	72	34	102	94	58	.79
	score ranks	3	6.5	4	6.5	1	2	5	
	prefer. ranks	3	5	2	7	1	4	6	
3	attitude score	132	82	112	74	142	134	82	.85
	score ranks	3	5.5	4	7	1	2	5.5	
	prefer. ranks	3	5	2	7	1	4	6	
4	attitude score	172	138	152	114	190	174	122	.86
	score ranks	3	5	4	7	1	2	6	
	prefer. ranks	3	5	2	7	1	4	6	
5	attitude score	207	180	187	149	225	209	157	.86
	score ranks	3	5	4	7	1	2	6	
	prefer. ranks	3	5	2	7	1	4	6	
6	attitude score	249	210	211	173	261	245	187	.89*
	score ranks	2	5	4	7	1	3	6	
	prefer. ranks	3	5	2	7	1	4	6	
7	attitude score	253	214	215	177	265	249	191	.89
	score ranks	2	5	4	7	1	3	6	
	prefer. ranks	3	5	2	7	1	4	6	

* $\rho_{\max} = .89$, at stage 6

Table 1 (continued)Respondent # 28

Model form: $A_j = \sum I_k^* \cdot B_{jk}^*$

Order of Attribute
Entry

2 3 1 4 6 5 7

Determinism score

1.66 .91 .69 -.06 -.35 -.54 -1.29

store number

<u>stage</u>		<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	<u>RHO</u>
1	attitude scores	.78	.78	-.13	-.13	1.70	-1.05	-1.96	.76
	score ranks	2.5	2.5	4.5	4.5	1	6	7	
	prefer. ranks	1	2	6	5	3	4	7	
2	attitude scores	1.65	1.65	-.48	-.48	1.36	-.18	-3.51	.98*
	score ranks	1.5	1.5	5.5	5.5	3	4	7	
	prefer. ranks	1	2	6	5	3	4	7	
3	attitude scores	2.03	2.42	-.48	-.48	1.36	-.18	-4.68	.96
	score ranks	2	1	5.5	5.5	3	4	7	
	prefer. ranks	1	2	6	5	3	4	7	
4	attitude scores	2.00	2.35	-.48	-.48	1.36	-.15	-4.61	.96
	score ranks	2	1	5.5	5.5	3	4	7	
	prefer. ranks	1	2	6	5	3	4	7	
5	attitude scores	1.89	2.62	-.20	-.20	.86	-.64	-4.30	.85
	score ranks	2	1	4.5	4.5	3	6	7	
	prefer. ranks	1	2	6	5	3	4	7	
6	attitude scores	1.39	2.13	.17	.17	.37	-1.14	-3.09	.85
	score ranks	2	1	4.5	4.5	3	6	7	
	prefer. ranks	1	2	6	5	3	4	7	
7	attitude scores	2.29	3.02	1.07	1.07	-.83	-4.42	-2.20	.56
	score ranks	2	1	3.5	3.5	5	7	6	
	prefer. ranks	1	2	6	5	3	4	7	

* ρ max = .98, at stage 2

Table 2

Results for the Traditional Approach
Spearman Rho Correlations
Between Attitude and Preference Ranks
(all 7 attributes included)

<u>respondent #</u>	<u>model I</u>	<u>model II</u>
1	.93	-.07
2	.90	.60
5	.77	.03
6	.07	-.63
8	.81	.14
9	.94	.49
11	.89	.82
14	.64	.53
16	.78	.96
17	.77	.66
18	-.30	.60
19	.94	.49
22	.64	.37
23	.97	.05
26	.31	.97
29	.95	.47
30	-.22	.07
31	.09	-.14
33	.56	.60
34	.21	-.10
35	.96	.55
37	.91	.06
38	.41	.38
39	.50	.90
41	.80	.33
42	.20	-.90
44	.50	.90
46	.68	.54
47	.97	.90
<u>mean*</u>	<u>.60</u>	<u>.36</u>

*computed after Fisher's Z-transformation.

Wilcoxon matched-pair signed-ranks test:
model I values significantly higher than
model II values ($p \leq .0001$)

Table 3

The Effects of the Number of Attributes
Included in the Model: Mean Correlations
At Each Entry Stage
(attribute sequence constrained to be
the same for all individuals)

	number of attributes included(stage)						
	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>
model I	.40	.54	.54	.58	.63	.60	.60
model II	.40	.54	.54	.59	.60	.51	.36

Figure 2

Mean Correlations Plotted Against the
Number of Attributes Entered

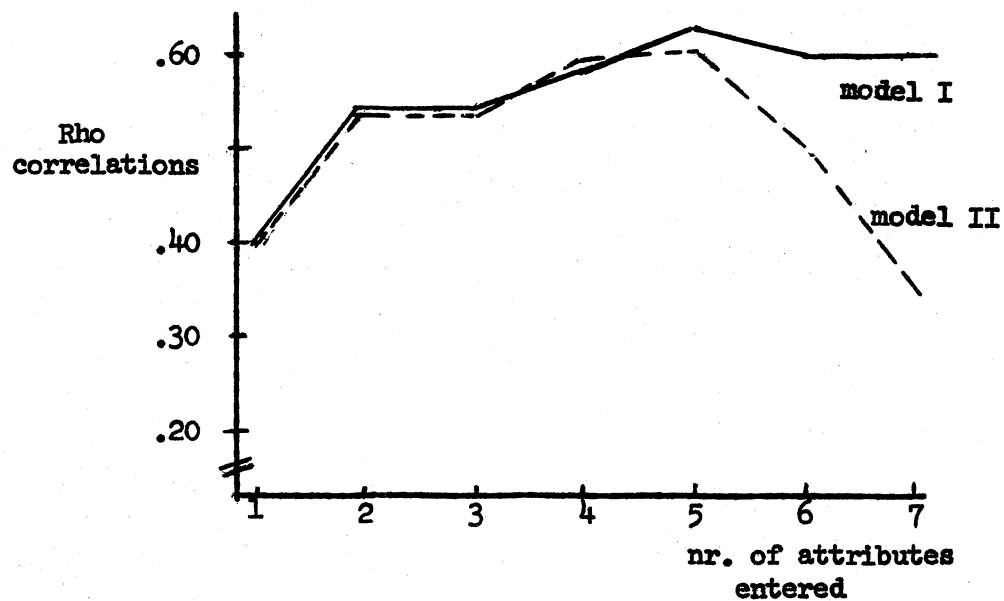


Table 4

The Effect of the Type of Attributes
Included in the Model: Mean Correlations
At Each Entry Stage
(attribute types permitted to vary at
each stage for each individual)

	number of attributes included						
	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>
model I	.50	.52	.62	.60	.61	.61	.60
model II	.50	.52	.61	.61	.57	.48	.36

Figure 3

Mean Correlations Plotted Against the
Number of Attributes Entered (attribute
types permitted to vary
for each individual)

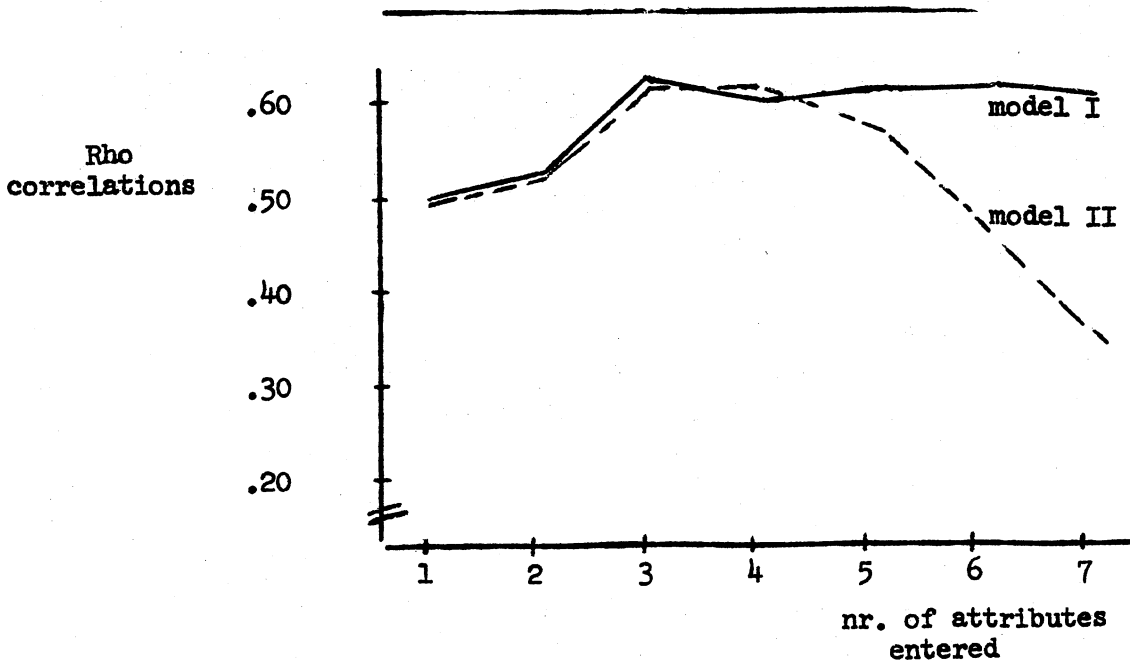


Table 5

Maximum or "Peak" Correlations (ρ max) vs
Correlations at Stage 7 (ρ 7)
(by respondent and model form)

respondent	model I		model II		model I	model II
	ρ max	at stage	ρ max	at stage	ρ 7	ρ 7
1	.93	7	.92	1	.93	-.07
2	.90	3	.90	3	.90	.60
5	.77	7	.62	1	.77	.03
6	.37	5	.04	1	.07	-.63
8	.84	1	.84	1	.81	.14
9	.94	6	.77	3	.94	.49
11	.89	6	.86	3	.89	.82
14	.64	6	.61	1	.64	.53
16	.81	2	.96	4	.78	.96
17	.93	1	.93	1	.77	.66
18	.87	1	.87	1	-.30	.60
19	.94	3	1.00	3	.94	.49
22	.82	3	.87	3	.64	.37
23	.97	4	.97	1	.97	.05
26	.87	2	.97	5	.31	.97
29	.95	5	.47	3	.95	.47
30	.06	1	.07	6	-.22	.07
31	.94	3	.77	3	.09	-.14
33	.56	5	.60	4	.56	.60
34	.21	1	.21	1	.21	-.10
35	.96	5	.96	3	.96	.55
37	.91	5	.81	4	.91	.06
38	.51	2	.51	2	.41	.38
39	.97	1	.98	1	.50	.90
41	.91	1	.91	1	.80	.33
42	.20	5	-.36	1	.20	-.90
44	.60	3	.90	3	.50	.90
46	.86	4	.93	4	.68	.54
47	.97	6	.90	4	.97	.90
means	.76*	3.6	.71	2.5	.60	.36

*column averages after Fisher's Z-transformation.

Wilcoxon matched-pairs signed-ranks test on magnitude differences: column 1 values statistically different from values, column 3; column 2 values statistically different from values, column 4 ($p \leq .001$)

Table 6

Model Performance at the Individual Level
 Number of respondents for which Fewer Attributed included ($k < 7$) resulted in Better, Equal and Worse Performance than using $k = 7$.

	<u>model I</u>	<u>model II</u>
$e_{\max} > e_7$	14 (48%)	22 (76%)
$e_{\max} = e_7$	13 (45%)	7 (24%)
$e_{\max} < e_7$	2 (7%)	0 (0%)

Table 7

Frequency Distribution of the Number of Attributes Included to Achieve e_{\max} (based on the "earliest" maxima for each respondent)

	number of attributes							
	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	<u>mean</u>
model I	7	3	5	2	6	4	2	3.6
model II	12	1	8	6	1	1	0	2.5

Table 8

Frequency Distribution of Attribute Types Included to Achieve e_{\max}

	<u>location</u>	<u>price</u>	<u>quality</u>	<u>variety</u>	<u>closeness to other services</u>	<u>shopping climate</u>	<u>other shoppers</u>
model I	19	23	20	18	14	7	2
model II	13	19	18	14	9	3	0

4. Finally, a fourth analysis allows for complete freedom in the number and type of attributes included for each individual. Using the individual determinism entry criterion, attitude scores and correlations for each respondent were computed each time an additional attribute was entered. As many ratings as necessary to achieve maximum correlation with preference were chosen for each respondent. Table 1 shows staging procedure results for two selected respondents to illustrate this procedure. It displays the entry order of attributes, the attitude scores computed at each stage, their ranks, and the Rho correlations with preference.

The Models

The analytical procedure is not straightforward for this problem. Results can depend upon various forms of the basic summative model used. Therefore, two (pragmatic) versions were included for analysis: (a) the traditional I·B model using raw scores (formula 2 - henceforth identified as model I), and (b) a more sensitive version, in which both the belief and importance scores are standardized before multiplication and summation:

$$(5) \quad A_j = \sum_k I_k^* \cdot B_{jk}^* \quad \text{model II}$$

where for each respondent,

A_j = the attitude score for store j

I_k^* = the standardized importance score for attribute k

B_{jk}^* = the standardized rating (across stores) of j on attribute k

Results

Study 1: The Traditional Approach

Table 2 shows the results of utilizing all seven ratings for each respondent (under the two model forms) and thus the results of a traditional multi-attribute analysis. For model I, the mean correlation (after Fisher's Z-transformation) was .60, with individual values ranging from -.30 to .97. For model II, the mean correlation was only .36, with a range of values between -.90 and .97. A Wilcoxon matched-pairs test indicated that the magnitude of correlation differences between the two model forms was statistically significant ($p \leq .001$).

Study 2: The Effect of the Number of Attributes

Table 3 and figure 2 summarize the results when the same attributes are included for every individual at each stage. The first attribute entered into the summation was #3 (quality). Thus, had only this dimension been considered for all respondents, a mean correlation of .40 would have been obtained. Similarly, choosing two common attributes for all individuals resulted in a mean correlation of .54.

Figure 2 plots mean correlations against number of attributes included, for both model forms. It shows that model performance increases asymptotically up to 5 attributes included. For model I, correlations at this level are not statistically different from using all seven attributes. This indicates that the original list could have been safely trimmed to 5 dimensions, without any loss in efficiency. For model II, however, the inclusion of all seven attributes yielded a considerably lower predictive power than the one possible with the selection of fewer attributes (4 or 5; attribute #3, 2, 4, 1, 5). The same

pattern for both model forms resulted when Importance scores were used as entry criterion instead of Determinism, so that poor performances seem to be model-dependent.³

Study 3: The Effect of Attribute Types

Table 4 and figure 3 show, in turn, the effect of varying the type of attributes included at each stage for each respondent. As will be recalled, idiosyncratic differences were allowed in the types of attributes chosen for each respondent, utilizing the individual Determinism scores to decide order of entry into the summation formula.

The analysis is more meaningful when the results are compared with those from study 2. For example, at stage 1 (one attribute included) the mean correlations under study 2 assumptions are .40 for both model forms versus .50 in study 3. This difference reflects the effect of utilizing one common attribute for all individuals versus the results of permitting its type to vary from respondent to respondent.

The impact of attribute types on performance is further evidenced by the earlier "peaks" obtained. Only 3 ratings were necessary to achieve maximum correlation when individual differences were allowed, as opposed to 5 in the previous approach.

Study 4: The Effect of Type and Number of Attributes

As explained earlier, the results of this analysis represent the most general approach, in that an optimal number and mix of attributes to achieve best fits is selected, independently for each respondent. Table 5 presents the maximum correlations (ρ max) for both model forms, as well as the number of attributes that were needed to achieve that "peak". Columns 3 and 4 are the results of using all seven attributes, as in the traditional approach (reproduced from Table 2 for comparison purposes).

Inspection of the results show that significantly higher correlations are achieved when individual differences are recognized (a mean of .76 versus .60 for model I; .71 versus .36 for model II). Also, the dispersion of values is reduced, compared to the results of the traditional approach.

Tables 6, 7, and 8 summarize the results at the individual level more succinctly. Table 6 shows, for example, that using the traditional I·B model form (model I) 48% of the respondents suffered from having all attributes included, while only 7% (2 respondents) gained. The remaining 45% did not change. More dramatic results are obtained when the standardized model form (model II) is utilized. Here, allowing for individual differences in the number of attributes included, performance increased for more than 76% of the sample.

Considerable individual differences were also found in terms of the order in which attributes should be included, and in the minimum number of attributes required for best fits. Table 7 presents these results. Utilizing model I, 24% achieved maximum rho with only one attribute included; 52% with three or less; while only 7% (2 respondents) with all attributes included. Similar observations also hold for the standardized model.

While Table 6 shows that the staging approach does improve the performance for a considerable number of individuals, it does not disclose the magnitude of these improvements. The Wilcoxon test for matched pairs (on data of Table 5) was used to test for magnitude differences in the correlations for the two model forms against the typical approach. In all cases, the differences were found to be statistically significant beyond the .01% level, thus confirming the hypothesis that the recognition of individual differences will improve over the typical approach.

Conclusions - Implications

It appears that the underpinnings of marketing's use of the multi-attribute attitude model call for explicit provision for individual differences in attitude dimensionalities, particularly given the acceptance of perceptual attributes rather than controllable product characteristics. Because of a relative lack of research on the issue, however, little empirical work has used this approach.

This study compared the use of all 7 available store attributes against staged attribute entries which allow for each respondent's brand attitudes to be represented by some subset of the original 7 attributes. It is clear that the latter approach offers predictions of store preference which are both more efficient and significantly higher. The following conclusions are thus advanced:

1. In accord with related results in choice theory and information processing, it appears that attitudes can be efficiently described (in predictive terms) with fewer attributes than are typically gathered in marketing research.
2. In accord with the theoretical proposals of Rosenberg and Fishbein, incorporation of only "salient" attributes (which may well differ by individual) leads to significantly better results than inclusion of all available ratings, and, following, that the typical practice of using all attributes for every respondent is likely to significantly understate the predictive power of the multi-attribute attitude model in marketing.

At the same time it should be noted that the more flexible approach advanced in the paper has its drawbacks. Cross-sectional analyses aimed at descriptions of variables (attributes) are now much more difficult; summarizing results in any form beyond predictive power is drastically hindered. Similarly, this approach makes cross-validation cumbersome, if not impossible. Methodological questions of the staging approach are by no means settled; "importance" and "determinism" served here as entry criteria, but it is likely that a better measure of "salience" can be developed. A weakness of standardization was apparent in these results. It is not clear how heavily results depend on the initial specification of attributes; this study accepted a list of seven as exhaustive.

It does seem, however, that the results of this study provide sufficient evidence to conclude that this aspect of the attitude model warrants further investigation. Future research should benefit from provisions for differences in the number and nature of attributes included in the modeling of each individual's brand attitude structure.

Footnotes

1. William L. Wilkie, Assistant Professor; and Rolf P. Weinreich, Doctoral Candidate, Krannert Graduate School, Purdue University.
2. The importance criterion simply used the magnitudes of respondent's stated importance weights to rank-order attributes for inclusion. Tied ratings were resolved by randomized rank assignments.
3. This interesting result is due to the impact of standardization on ratings which have little disparity across stores, such that one store which diverges slightly in either direction is stretched further and often shifts several places in attitude score ratings. Little disparity appears much more frequently in unimportant attributes.

References

- Anderson, L.R. & Fishbein, M. Prediction of attitude from the number, strength, and evaluative aspect of beliefs about the attitude object: a comparison of summation and congruity theories. In M. Fishbein (Ed.), Readings in attitude theory and measurement. New York: Wiley and Sons, 437-443.
- Fishbein, M. A behavior theory approach to the relations between beliefs about an object. In M. Fishbein (Ed.), Readings in attitude theory and measurement. New York: Wiley and Sons, 389-399.
- Fishbein, M. A consideration of beliefs, and their role in attitude measurement. In M. Fishbein (Ed.), Readings in attitude theory and measurement. New York: Wiley and Sons, 257-266.
- Hansen, F. Consumer choice behavior: an experimental approach. Journal of Marketing Research, 1969, 6, 436-443.
- Myers, J.H. & Alpert, M.I. Determinant buying attitudes: meaning and measurement. Journal of Marketing, 1968, 82, 13-20.
- Osgood, C.E., Suci, G.J., & Tannenbaum, P.H. The measurement of meaning. University of Illinois Press, 1957.
- Rosenberg, M.J. Cognitive structure and attitudinal affect. In M. Fishbein (Ed.), Readings in attitude theory and measurement. New York: Wiley and Sons, 325-331.
- Wilkie, W.L. Issues in marketing's use of multi-attribute attitude models. Purdue University, Krannert Graduate School. Institute paper, 1972.

AN APPROACH TO THE RESOLUTION OF MULTICOLINEARITY
IN THE ATTRIBUTE STRUCTURE OF ATTITUDES¹

Reza Moinpour
University of Washington
and James B. Wiley²
University of British Columbia

Recently, a number of market researchers have proposed that the consumer's preference for a product can be accounted for by a weighted, additive utility (WAU) model (Bass and Talarzyk, 1969; Cohen and Ahtola, 1971; Cohen and Houston, 1971; Hughes, 1970; Moinpour and MacLachlan, 1971; Sheth, 1970; Sheth and Talarzyk, 1970). The WAU model is derived from attitude theories formulated by Rosenberg (1956) and Fishbein (1965).

Such models postulate that the individual's preference for a product is a function of a) the degree to which the product possesses certain attributes, and b) the importance of the attributes. One example of such a construct is the following (Moinpour and MacLachlan, 1971):

$$(1) \quad A_x = \sum_{i=1}^n W_i B_{ix}$$

where A_x = a consumer's attitude toward a particular product or brand X .

W_i = the importance or weight of attribute i .

B_{ix} = the product's satisfaction score on attribute i ; subject's belief about attribute i for product X .

n = the number of product attributes.

A review of the literature reveals some important issues which market researchers have encountered in applications of WAU model in consumer behavior:

1. The use of separate or disaggregative rather than the summed-score form of the model.
2. The application of individual vs. group analysis as methodological procedures.
3. The extent of the contributions of the two components of the model (W_i and B_{ix}) to its predictive power.
4. The existence of collinearity amongst individual's evaluative beliefs.

There appears to be agreement among researchers that the disaggregative form of the model is the more appropriate one for the investigation of consumer attitude (Cohen and Ahtola, 1971; Cohen and Houston, 1971; Sheth, 1970). In this form, the model allows for further evaluation of the relative contributions of underlying attributes.

Most of the studies in this area have used regression as their method of analysis. Recently, applications of multidimensional scaling (Hansen and Bolland,

1971; Moinpour and MacLachlan, 1971) and canonical analysis (Lutz and Howard, 1971; Sheth, 1971) have also been reported. However, the generalization across subjects that takes place when a group is used as the experimental unit may lead to confounded response circumstances. This problem can be attacked by carrying out regression on groups of subjects with homogeneity of response sets (Scott and Bennett, 1971). It has been correctly pointed out, however, that the theory underlying the WAU model depicts an individual construct; thus, individual analysis is conceptually the proper one to use (Moinpour and MacLachlan, 1971).

There has been some controversy surrounding the question of the extent of the contributions of the two components of the attitude model. Several researchers have suggested that "weights" incorporated into the WAU model contribute little to its predictive power (Cohen and Houston, 1971; Moinpour and MacLachlan, 1971; Sheth and Talarzyk, 1970). The unimpressiveness of respondent-provided "weights" to the model's predictive power has been a factor in the decision of a number of researchers to use regression analysis to estimate weights rather than use "weights" supplied by respondents (Cohen and Ahtola, 1971; Sheth, 1970). We argue that, given the theory underlying the model (1), the lack of contribution of weights to the predictive power may be due to the fact that because of measurement problems the attribute scores may in fact already reflect the "weighting" criterion. In other words, they may in fact be "weighted judgments;" consequently, multiplying by independently derived "weights" would serve only to doubly weight the attribute scores. It has been suggested, as an alternative approach, that the attribute importance component of the model can best serve as criteria for the selection of attributes (Moinpour and MacLachlan, 1971).

A final problem, which infrequently has been considered, is the existence of intercorrelation between individuals' expressed, evaluative beliefs concerning brands' possession of "salient" product attributes. One might expect this problem to arise in the case of unplanned data. However, the results of a number of studies using factor analysis (Moinpour and MacLachlan, 1971) and canonical correlation (Sheth, 1971) suggest that the phenomenon may occur even when care is taken in the original selection of "salient" product attributes. Inasmuch as the model (1) prescribes an independent structure, we feel that an effort should be made to satisfy this requirement. This is particularly true when regression techniques are used, since estimates of regression weights are notoriously unstable when collinearity exists among the explanatory variables (Kendall, 1957).

The Objective

This study is concerned with the resolution of the collinearity problem in the attribute structure. In this paper, we describe a method for selecting a subset of "salient" product attributes which minimizes overlapping of information supplied by explanatory variables in regressions. Factor analysis is employed to summarize a set of potential explanatory variables to one with a few variables with a little loss of information. It can be used to identify and to extract a relatively uncorrelated subset of independent variables to be included in regressions. In this study, predictions of brand preferences (for soft drinks) made from an uncorrelated subset of three product attributes were often better, and never significantly worse than, predictions made using the entire set of ten product attributes. Furthermore, in light of previous findings, it was encouraging that those attributes selected for the model happened also to be those indicated by the respondents to be "important."

The Data

Data was collected from 35 University of Washington students. The following information for 10 brands of soft drinks were collected:

Part 1. Each student was asked to rate each brand on 10 attributes using a 6-point scale, ranging from 1--satisfactory to 6--unsatisfactory. The list of brands and product attributes appears in Table 1.

Part 2. Each respondent was asked to evaluate these product attributes in terms of importance on a 6-point scale, ranging from 1--important to 6--unimportant.

Part 3. The participant was also asked to rate the 10 brands in terms of preference using a 10-point scale, ranging from 1--most prefer to 10--least prefer.

Table 1

List of Brands and Attributes

Brands	Attributes
Pepsi Cola	Color
Coca Cola	Taste
Tab	Carbonation
Fresca	Price
7-Up	Calories
Dr. Pepper	Sweetness
Hires root beer	Thirst Quenching
Hi-C, grape	Liking of flavor
Hawaiian Punch	After-taste
Orange Crush	Packaging

Method of Analysis

The objective is to minimize the intercorrelation between explanatory variables in a regression equation. One method suggested by Kendall (1957) is to regress the dependent variable (preference) on the principle components of R_{ii} , the correlation matrix of the total pool of explanatory variables (product attributes). The result, by definition, is orthogonal regression. We adopt an approach based on the one proposed by Kendall (1957) and previously utilized by Twedt (1952) and Daling and Tamura (1970). The aim of this latter approach is to select a set of relatively uncorrelated variables, using the extracted factors as guidelines, that retain most of the predictive power of the original pool of potential explanatory variables. The basic factor analysis model is of the following form:

$$(2) \quad X_i = \sum_{j=1}^n a_{ij} f_j + e_i$$

where X_i = value of an observed explanatory variable

f_j = common factors

a_{ij} = factor loadings, indicates the correlation between X_i and f_j

e_i = an error term

The principle factor loading matrix, A , for the first three principle components of R_{ij} was calculated and rotated according to the varimax criterion. An element of this matrix, a_{ij} , indicates the correlation between attribute X_i and factor f_j . The varimax rotation tends to maximize the correlation between an attribute and a single factor. If an attribute is highly correlated with a factor and the factor is highly correlated with preference, then, in turn, the attribute should also be correlated with preference. Similarly, if an attribute is highly correlated with a factor it cannot be highly correlated with another factor and, in turn, cannot be highly correlated with attributes having high loadings on that factor. Thus, by selecting attributes having maximum loadings on the rotated factors, a set of relatively uncorrelated explanatory variables were obtained.

The success of this procedure can be evaluated by performing two sets of regressions for each product (across subjects). One regression incorporates the entire pool of potential explanatory variables (product attributes), the other includes only the three attributes selected in the manner described above. That is, for each brand we fit the following model:

$$(3) \quad Y = a + \sum_{i=1}^n b_i X_i + e$$

where Y = stated preference ratings

X_i = the attribute scores

e = an error term

The results appear in Table 2.

Discussion of Results

For each of the ten brands separate regressions were performed of preference rating on attribute scores across subjects, using all attributes. Next, the subject by attribute matrix of each of the previous regressions was factor analyzed. The first three factors extracted accounted for most of the variance of the explanatory variables. Following a procedure adopted by Twedt (1952) and by Daling and Tamura (1970), the product attributes with the highest loading on each of the rotated factors were selected for inclusion in a second set of regressions. These variables offer most promise for prediction of preference.

Adjusted R^2 values for both sets of regressions are given in Table 2. (Adjusted coefficients of determination are presented to allow comparison with other studies.) For six of the ten brands, the reduced attribute set explained

Table 2
Summary of Regression Results

Brands	All Variables	Selected Variables	Variables Selected in the Prediction Equation
Pepsi Cola	.4279**	.3703**	Liking of flavor, color, price
Coca Cola	.4559**	.3796**	After-taste, color, price
Tab	.6041**	.6375**	Taste, sweetness, calories
Fresca	.6253**	.6915**	Taste, carbonation, price
7-Up	.2843*	.1880*	Taste, thirst quenching, sweetness
Dr. Pepper	.4656**	.4987**	Taste, color, price
Hires root beer	-0	.0684	Taste, calories, packaging
Hi-C, grape	.5033**	.5041**	Liking of flavor, calories, taste
Hawaiian Punch	.4087**	.3695**	Liking of flavor, taste, carbonation
Orange Crush	.4475**	.5051**	Taste, liking of flavor, after-taste

Note: "-0" indicates an R^2 value so close to zero that the corresponding adjusted R^2 is negative.

*p < .05

**p < .01

more of the variation in the preference ratings than did the complete attribute set. For the remaining four brands the results of the two regressions were comparable. We infer from these results that reduction of variables in the regression does not lead to significant reduction in R^2 when the reduced variable set is free from the interdependency which may have existed in the original variable set. We may add that most of the variables selected on the basis of the factor analysis were indicated, a priori, by the respondents as being important product attributes.

Concluding Remarks

It appears that the major issues confronting the researchers in this area relate to the specification of the variables and their measurement and the specification of the WAU model itself. A critical aspect of the first problem is the existence of intercorrelation among variables even when care is taken in the original selection of "salient" product attributes (Lehman, 1971). Since the model (1) prescribes an independent structure, attempt should be made to satisfy this requirement especially when regression is used since the net regression coefficients tend to be unstable when collinearity exists among the explanatory variables (Kendall, 1957). An approach to this problem is the use of orthogonal factor analysis for selection of variables. This technique provides the experimenter with an orthogonal subset of explanatory variables to be used in regressions. The results of this study indicate that significant prediction of brand preference can be made from an uncorrelated subset of explanatory variables (product attributes).

Footnotes

1. The research was supported by the Office of Publications and Research, Graduate School of Business Administration, University of Washington.
2. Reza Moinpour is Assistant Professor of Marketing, Graduate School of Business Administration, University of Washington, Seattle, Washington. James B. Wiley is Acting Assistant Professor of Marketing, Faculty of Commerce and Business Administration, The University of British Columbia, Vancouver, British Columbia, Canada.

References

- Bass, F. M. & Talarzyk, W. W. A study of attitude theory and brand preference. Krannert Graduate School of Industrial Administration, Purdue University, 1969.
- Cohen, J. B. & Ahtola, O. T. An expectancy X value analysis of the relationship between consumer attitude and behavior. Proceedings. 2nd Annual Conference, Association for Consumer Research, 1971, 344-364.
- Cohen, J. B. & Houston, M. J. The structure of consumer attitudes: the use of attribute possession and importance scores. University of Illinois, College of Commerce and Business Administration, Working paper no. 2, January 1971.
- Daling, J. R. & Tamura, H. Use of orthogonal factors for selection of variables in a regression equation--an illustration. The Journal of the Royal Statistical Society, 1970, 19, 260-268.
- Fishbein, M. A consideration of beliefs, attitudes, and their relationships. In I. D. Steiner and M. A. Fishbein (Eds.), Current studies in social psychology. New York: Holt, Rinehart and Winston, 1965.
- Hansen, F. & Bolland, T. The relationship between cognitive models of choice and non-metric multidimensional scaling. Proceedings. 2nd Annual Conference, Association for Consumer Research, 1971, 376-388.
- Hughes, G. D. Distinguishing salience and valence. Paper presented at the Attitude Research and Consumer Behavior Workshop, University of Illinois, December 1970.
- Kendall, M. G. A course in multivariate analysis. New York: Hafner Publishing Co., 1957.
- Lehman, D. R. Television show preference: application of a choice model. Journal of Marketing Research, 1971, 8, 47-55.
- Lutz, R. J. & Howard, J. A. Toward a comprehensive view of the attitude-behavior relationship: the use of multiple-set canonical analysis. Proceedings. Social Statistics Section, American Statistical Association, 1971, 215-225.
- Moinpour, R. & MacLachlan, D. L. Relations among attribute and importance components in Rosenberg-Fishbein type attitude model: an empirical investigation. Proceedings. 2nd Annual Conference, Association for Consumer Research, 1971, 365-375.
- Rosenberg, M. J. Cognitive structure and attitudinal affect. Journal of Abnormal and Social Psychology, 1956, 53, 367-372.
- Scott, J. E. & Bennett, P. D. Cognitive models of attitude structure: 'value importance' is important. Paper presented at the Fall Conference of the American Marketing Association, Minneapolis, 1971.

Sheth, J. N. An investigation of relationships among evaluative beliefs, affect, behavioral intention and behavior. Mimeographed paper, April 1970.

Sheth, J. N. Canonical analysis of attitude-behavior relationship. Mimeographed paper, March 1971.

Sheth, J. N. & Talarzyk, W. W. Relative contribution of perceived instrumentality and value importance component in determining attitudes. Paper presented at the Fall Conference of the American Marketing Association, Boston, 1970.

Twedt, D. W. A multiple factor analysis of advertising readership. Journal of Applied Psychology, 1952, 34, 207-215.

WARRANTY POLICIES AND PRACTICES
OF CONSUMER PACKAGED GOODS MANUFACTURERS

C. L. Kendall
Frederick A. Russ
University of North Carolina

One manifestation of our times is the increasing public demand for more protection, performance, and responsiveness from its institutions. In a high consumption society it is inevitable that these demands will fall heavily on the consumption artifacts of that society and the organizations that make them. Consumerism is no recent phenomenon however. The current furor over alleged product failure and maker responsibility is only the latest peak in a long controversy over the determination of obligation and responsibility in the buyer-seller transaction. The present swell of consumerism is, however, distinguished by its intensity and breadth--probably due to such factors as mass communications broadly consumed by an increasingly well-educated and aware public.

A related, concurrent development has been the increased tempo of change in the legal relationships between buyer and seller caused by judicial recognition of the changes in commerce created by mass production and distribution. The retailer-consumer relationship has long been in decline: self service spreads and the large national manufacturer communicates directly with the consumer through national advertising. The old legal relationship between the consumer and the manufacturer was two or three times removed. Under the law of contract the consumer who sustained injury had recourse only to the retailer who sold him the goods; the retailer to the wholesaler; and so back through the channel to the manufacturer--an expensive, time-consuming, and cumbersome process of litigation. In recent years a direct consumer-manufacturer relationship is increasingly recognized by the courts: an injured or disappointed consumer may be able to seek direct legal redress against the manufacturer.

Although most legal actions against manufacturers are still based on injury through negligence, product liability arising from warranty is becoming more and more important. Under warranty, fault or failure--i.e., negligence on the part of the manufacturer--is irrelevant. The only thing that need be proved is that a warranty existed, and that there was some breach of it.

There are two types of warranty. The first is an express warranty by the manufacturer that promises or disclaims something, e.g., "replacement of (certain) defective parts for no more than one year from the date of purchase." Much confusion exists as to whether, and how much, sales puffing constitutes a warranty. The second type of warranty is an implied warranty--simply that the goods sold will be appropriate for, and perform effectively in the commonly understood function of that good. This warranty exists in all transactions unless superseded or somehow modified by an express warranty. Aside from the disclaimer or obligation-limiting aspect of warranties, they often have a promotional nature as well--e.g., the ubiquitous "money back" guarantee--and true product-differentiating features such as the guarantee on the 1972 American Motors automobiles.

In recent years much attention has been focused on the warranties of automobiles and other major consumer expenditure items, principally appliances. The literature on this topic deals primarily with the capability, willingness, and

alacrity of the manufacturer in making good on the claims of the buyer against the explicit specifications in the purchase contract. Some attention has also been paid to the rights of claimants under the doctrine of implied warranty as well. The attention paid warranties of these major consumer expenditure items is natural; they represent major investments and high potential for serious inconvenience in the case of malfunction.

Warranties are also given or implied by promotion or package copy for "small" consumer expenditure items such as packaged foods, health and beauty aids, common home-use products, etc. These goods, although individually less important than "big ticket" items, account for a large proportion of total consumer expenditures and an even larger proportion of total transactions. In spite of this, a search of the literature suggests that little is known of the warranty practices of consumer packaged goods manufacturers (CPGMs).

A malfunction or a frustrated expectation in a product costing \$.89 or \$1.49 is far more likely to result in a muttered expletive and a subsequent brand switch than in some kind of complaint. Recourse to litigation by an individual consumer for such a product "failure" (unless it is botulism in vichy-ssoise) is almost beyond the realm of comprehension. Even a complaint to the manufacturer is unlikely when a low priced packaged item fails: the trade-off between consumer effort and the possibility of a satisfactory corporate response--apology, refund, or replacement--is heavily weighted in favor of taking no action.

An area of such size--in total consumer expenditure and transactions and in potential abuse and ill will to the business system--deserves some illumination by more information. The remainder of this paper is devoted to describing the results of exploratory research designed to shed some light on CPGM warranty practices and policies. Unlike automobile and appliance manufacturers, CPGMs sometimes guarantee both quality and satisfaction. The latter is undoubtedly possible and a practice because of the previously-mentioned inhibitions to complaints or claims, and the inexpensive nature of a refund for those few complaints or claims that are made under such a warranty. A search of the literature suggests that little information exists on (1) which manufacturers or industries give such warranties, (2) how manufacturers respond to claims against warranties, (3) whether there are response differences between manufacturers with express warranties and those with implied warranties, (4) how manufacturers themselves view warranties and claims, and (5) how consumers perceive these warranties. Two studies were undertaken to explore the first three topics. Research on the last two is underway.

Shelf Audit Study

The objective of the first phase of the research was to determine what manufacturers gave express warranties, and the nature of these warranties. The methodology was a simple shelf audit of several hundred consumer packaged goods in a number of selected packaged goods categories. The survey was limited to package copy. It was felt that warranties would occur more frequently in package copy than in media advertising for that product, although this is a topic for future research.

The principal finding of this stage of the study was that most package goods of the type found in super markets and drugstores do not carry express warranties. Further, though the incidence of warranties is small in all categories, it does vary somewhat across product categories. For example, in all types of frozen foods surveyed (39 brands comprising roughly 200 different products), only two products, both from the same company, carried an express warranty, and only two others carried a Good Housekeeping Seal. In contrast, guarantees were common

in margarines - about one third. Most package copy for foods is limited to the mandatory listing of ingredients and an occasional puff concerning ingredients or process, e.g., "creamy", "selected", etc. or a statement such as "Superior taste with lighter texture". Such claims are unlikely to be interpreted as express warranties.

Overall, the incidence of promotional claims and warranties on food products is much less than in such product categories as health and beauty aids and automotive and home use products. Warranties could provide a basis for product differentiation, because in no industry studied were all major brands guaranteed. But warranties don't seem to be used in product differentiation strategies, because it was rare to find the guarantee featured prominently on the package.

Warranty policy is anything but consistent. A manufacturer may have a money back guarantee on one product line and not on another; one product within a product line and not on the others; and, capriciously enough, in the case of two products discovered, a guarantee on one size of a product and not on the other size of the same product! In one major packaged goods industry, one of the three large companies that dominate the industry carried a satisfaction-or-money-back guarantee on all of its many products, while its two competitors had guarantees on none.

Almost all express warranties exhibited some variation of the "satisfaction guaranteed or money back" form rather than providing a list of legalistic disclaimers as do major expenditure items. Often the consumer packaged goods warranties made impossible or inconvenient demands on the dissatisfied consumer. To make a "legitimate" claim, consumers were required to send a sales slip, which is rarely saved, or the unused portion, which is costly and inconvenient to return--or may not even be allowed through the U.S. mail.

The low incidence of warranties and the apparent lack of a coherent warranty policy lead to several possible conclusions, all of which require additional substantiation.

1. CPGM's perceive little pressure from consumers or competitors to offer express warranties.
2. Where warranties are offered, product characteristics and company size may be determining factors. Those products with high social or technical risk (e.g., health and beauty aids, automotive additives) exhibit more guarantees than those with low social and technical risk. Furthermore, those companies which are small or not well-known are more likely to offer guarantees than large, well-known firms.

Corporate Response Study

The results of the shelf audits lead to several questions. If few CPGMs guarantee their products, do they still stand behind them? Do CPGMs with warranties respond "better" than those without? What forms do the responses of CPGMs to claims most frequently take?

The news media and informal research suggest a pronounced skepticism in America about CPGMs' response to claims. Many seem to feel that CPGMs are not concerned about consumer complaints (as evidenced, for example, by the lack of express warranties). Furthermore, CPGMs are not expected to make it easy for the consumer to get satisfaction: if CPGMs do respond, it is felt that they are likely to throw up roadblocks to hinder a consumer making a claim.

To find out whether these biases are justified, and to get answers to the previously asked questions, an exploratory study was conducted. Methodology and results of the study are described below.

Methodology

A judgment sample of approximately 200 different brands of consumer packaged goods was chosen to provide the initial data collection base for the study. The sample was reduced on the basis of the following decision rules.

1. Only those products with which dissatisfaction (rather than a product malfunction) could possibly exist should be considered.
2. Brands should be chosen from a wide variety of product categories, but, where possible, there should be several brands chosen from a given product category.
3. In each product category there should be an approximately even split between brand with warranties and those without. In product categories where warranties are especially scarce, a balance should be struck between those with strong package claims and those without.
4. Only one product should be chosen per manufacturer. (This limitation was forced by the apprehension that CPGMs might be alerted to the study if they received more than one complaint, almost identically phrased, from the same location.)

The resulting sample of 40 contained five brands in each of the following categories: insecticides, automotive oil additives, ready-to-eat cereals, shampoos, and deodorants. The remaining fifteen brands were widely scattered across mostly non-food product categories. Each of the brands chosen was nationally advertised and ranged in price from \$.19 to \$3.00 (median = mode = \$.99). Ten brands exhibited express warranties; the remainder made package claims of widely varying strength.

Each of the products chosen for the sample was purchased at a local super market or drugstore. Then, the products were tested to determine possible or actual reasons for dissatisfaction. After using the products, the researchers developed a standardized letter of complaint to be sent to each "participating" CPGM. The forty letters were identical except for handwriting and a specific complaint related to product performance and package claims. Within a given product category, letters were, as much as possible, identical. When a guarantee indicated that some proof of purchase was required, a sales slip, price tag, label, or boxtop was remitted with the letter.

Each letter was mailed--on the same day--to the manufacturer listed on the package. Although standard corporate directories were available to the researchers, it was felt that such a resource might not be readily available to the typical consumer; so, only the addresses printed on the product or package were used. When (if) responses to the complaint letters were received, the postmark date was duly noted and the responses were saved and categorized. The following information was collected:

1. Incidence of response
2. Speed of response (measured by the difference between the postmarked date of response and the original mailing date. If the response involved more than one piece of mail, then the earliest postmark date was used.)
3. Type of response (type of letter, method of providing satisfaction, position of respondent, etc.)

Results

Of the 40 letters of complaint which were mailed, 33 received some kind of response. There was no response (at least none within 2 months) to six of the letters, and one letter was returned due to an insufficient address. This last case is properly included with the no response category, because the dissatisfied consumer would not be easily able to reach that manufacturer to make a complaint.

Among those CPGMs who responded, the speed of response varied greatly. The average speed was 14.6 days, but eight CPGMs responded in either 7 or 8 days and another eight took more than three weeks. All who did respond within the two months allotted for the study did so within 4 weeks.

The response rate of 80+ percent can be caused either for optimism or for pessimism about the concern of CPGMs, depending on one's initial bias. Even an optimistic view must be tempered by recognition that not all of the responses would have been satisfactory to the complaining consumer. Rather than providing a refund or a replacement product, seven of the 33 responding firms wrote to ask for (additional) proofs of purchase or more information, or made some other temporizing response. The net effect of these "delaying" tactics (costly both to the customer and to the CPGM) was that the satisfactory response rate was reduced to only 65 percent of all letters sent.

Although the method by which the sample was drawn precludes tests of statistical significance, it is obvious that, in our sample, response for products with express warranties was essentially the same as for those without. Nine of the ten guaranteed products received a response (seven of those were satisfactory); 24 of the 30 non-guaranteed products received a response, 19 of which were satisfactory.

The 26 CPGMs who responded satisfactorily to consumer complaints took different approaches to satisfying those complaints. The most common response was a refund: 16 CPGMs sent a refund for the purchase price of the product, although none of them provided any monetary compensation for the consumer's cost of making the complaint. Three CPGMs did, however, send a replacement product as well as a refund. Another eight firms mailed either a replacement product or, in three cases, another product which, it was hoped, would be more satisfactory. The cost of mailing these products often exceeded their retail value; so it was surprising that only two CPGMs responded with coupons which entitled the consumer to a free product at his local retail store.

The response of firms was not limited only to refunds or replacement of the product. A number of CPGMs showed concern for the quality of the product and the conditions under which it was used: six firms requested additional information about the circumstances of use or purchase and another seven sent instructional material to help the consumer make better use of their products the next time.

The typical respondent to a complaint came from the customer service department of the firm (15 of the 27 whose positions were identifiable). Eleven of the remaining 12 identifiable respondents were in positions of such authority that they would be perceived to be able to "do something" about rectifying the cause of the complaint. Although most responses to complaints are probably "robotyped," only four of the responding CPGMs sent an obvious form letter. The personal or robotyped responses ranged from one short paragraph to three pages.

The range of response to consumer complaints may be illustrated by looking at the automotive oil additives product category. Letters were sent to 5 firms, 3 of whom provided express warranties for their products. All products retailed for \$.99. The manufacturer of product A (CPGM-A), with no guarantee, did not respond at all. CPGM-B, also without a guarantee, sent a letter and followed it up with a phone call and visit from a regional sales representative to see if the problem could be determined and solved. CPGM-C, with a guarantee that read "satisfaction guaranteed or your money back", sent a long, involved form which the consumer was requested to complete in order to receive a refund. And the manufacturer required a proof of purchase even though the guarantee did not state that such was necessary. (How often would you save the sales slip or the empty can for some oil additive you'd poured into your car's crankcase?) CPGM-D, with a double-your-money-back guarantee, also requested proof of purchase (a requirement unstated on the package) but at least made the request in a pleasant personal letter and did not require the completion of a tedious form. Finally, CPGM-E, also with a guarantee, sent a personal letter, a refund for the purchase price, and a lubrication and engine performance booklet to help the consumer understand the kinds of problems that the product could or couldn't cure.

Conclusions

The original bias indicated by current consumerist skepticism was not altogether justified. The consumer can get help when a product performs unsatisfactorily, even though the product may not be expressly guaranteed. CPGMs sometimes even go the extra mile: (1) they may provide a response over and above a refund of purchase price or replacement of the product and (2) they respond even though they may not be responsible for product failure. The essentially equivalent response for products with or without express warranties may mean that the lack of warranties on most consumer package goods is, or should be, of little importance to consumer decision-making.

The silver lining is not without its cloud, however. Some companies did not respond or, if they did respond, they made a temporizing response which caused additional effort to be expended by the consumer. The reasons for these practices are difficult to understand, particularly when the consumer movement is paying such close attention to the activities of all kinds of manufacturers. Equally puzzling is the fact that a few of the CPGMs failed to use their response as an opportunity to improve public relations and to avoid negative word-of-mouth advertising. Indeed, two of the worst "offenders" were companies which have been in recent public view because of charges levelled at them by the FTC about their promotional practices!

The big picture, if any clear picture emerges, is one of fragmentation or diversity. Some CPGMs guarantee their products; others do not. Some respond to consumer complaints; others do not. To understand the reasons for this diversity, additional research must be undertaken. A brief description of such research is provided in the remaining section of the paper.

Where-Do-We-Go-From-Here?

Two areas for future research suggest themselves. The first area focuses on what the consumer packaged goods industry is doing and why. A study is currently underway, using a mail survey approach, to obtain information about CPGM warranty policies and their response to consumer complaints. This survey will obtain information to answer questions such as: How many CPGMs guarantee

their products? Why? Who sets guarantee policies? Where do the guarantees appear? What do they say? Why do most CPGMs fail to guarantee their products? What is the volume of consumer complaints? Who responds to consumer complaints and how? What impact have complaints had on company decisions?

This information should be of value to CPGMs themselves, because the product guarantee is a part of the marketing mix. It may also be illuminating to consumer advocates, who want to know whether and how "strongly" CPGMs stand behind their products performance. [Preliminary results of this study will be available by the date of the ACR meetings.]

A second area of research, also of interest to CPGMs and to consumer advocates, is to investigate the behavior of consumers with regard to guarantees for consumer package goods. Such questions as the following need to be answered. Do consumers perceive guarantees as influential factors in purchase decisions? What types of consumers complain about products? Is the probability of a complaint increased by the presence of an express warranty? Are complaining consumers generally satisfied with corporate response to their complaints?

Information from manufacturers and consumers, along with the legal opinions about warranties which have already been rendered, should provide a balanced picture of warranty and claims policies and practices and indicate whether and where action by manufacturers, governments, and individual or groups of consumers would be appropriate.

THE LEGAL VERSUS THE BEHAVIORAL MEANING
OF DECEPTION

James E. Haefner
University of Illinois

Introduction

The legal meaning of deception has been developed through numerous Federal Trade Commission and court decisions. For the most part, these decisions have not relied upon consumer evidence. In some instances, this appears to be a justifiable position. A claim that the product "Is Made of Cotton" (when it in fact is not) is so obvious that consumer testimony would not be useful. Claims dealing with technical features of a product ("Brand X Tire will last 30,000 miles under normal driving conditions," or "X Air Conditioner completely cleans the air of all dust and foreign particles") would also not require consumer data as only careful laboratory tests could determine their truthfulness.

There are other situations, however, where consumer data is needed. Direct evidence is needed when the interpretation advanced depends on specific assumptions about consumer behavior. Without this direct evidence, the legal meaning of deception must be suspect.

The "doctrine of tendency to deceive" has been used by the FTC and the courts to exclude consumer evidence whether specific assumptions about consumer behavior are needed or not (Kinter, 1971). This doctrine states that actual consumer deception need not be proven, only the tendency to deceive need exist. In order to prove that a tendency to deceive exists in an ad, the FTC has substituted other types of evidence for consumer data.

One method used by the Commission to replace consumer evidence is called the hunch, or intuitive, approach (the Commission knows from past experience and expertise how consumers will interpret advertising messages) (Gellhorn, 1969, p. 564). An example will help to point out its weakness. In FTC v. Mary Carter Paint Co.,² Mary Carter was selling paint for \$6.98 a gallon with a second gallon free. This promotion pricing strategy was used by the company because it felt that if it would sell its paint below the price of other national competitors, consumers, who judge the quality of paint by its price, would view their product as inferior. The Commission determined that this offer was deceptive because it implied that the price used was the usual and customary one for the single article. This conclusion is not obvious and calls for direct customer evidence.³

Another approach used in place of consumer testimony utilizes dictionary definitions (Gellhorn, 1969, p. 565). Again an example will point out its weakness. The fact that rejuvenescence means "a renewing of youth" tells little about whether consumers would give that same meaning to a word in an advertisement for a cosmetic called "Rejuvenescence Cream."⁴

A third method used to replace consumer data utilizes trade experts (Gellhorn, 1969, p. 565). Trade experts are called to the witness stand in order to relate their interpretation of what consumers perceive an ad or claim to mean. These experts may have a sophisticated view on the subject, but they do not necessarily have any expertise or information about actual consumer behavior concerning the issue in question.

The replacement of consumer data with these other approaches has lead to the following legal meaning of deception:⁵

An act or practice or representation is a "deceptive" act or practice, which Section 5 of the Federal Trade Commission Act prohibits, if it has the "capacity" or "tendency" to deceive.⁶ Actual deception need not be proved or found.

In determining whether a representation is deceptive, its effect upon the "ordinary" purchaser, the "public" or the "average" man is considered.⁸ This may include "the ignorant, the unthinking and the credulous"⁹ and "the least sophisticated."¹⁰ However, it does not include a "very stupid person"¹¹ or "an insignificant and unrepresentative"¹² segment of the public.

In determining whether a representation is deceptive, the entire representation is taken in consideration¹³ --the entire advertisement. If a word or term is ambiguous,¹⁴ and one meaning is false, the word or term is held to be deceptive.

Not only what is said, but what is not said may constitute a deceptive act.¹⁵ Representations which are too broad to be true in all circumstances must be qualified.

An act which may be deceptive to ultimate consumers or the general public is unlawful even though the immediate customer is not deceived.¹⁶

Although the above legal definition may have meaning for the adjudication process, it possesses no meaning in a behavioral sense. Since the FTC and the courts are dealing with a behavioral construct, that of deception, it is necessary to interpret this construct from the consumer's point of view.

In sum, the exclusion of consumer data and its replacement with evidence based on the hunch, the dictionary definition, or the trade-expert cannot yield an accurate meaning of deception. Deception is something that is perceived by the consumer, and it is only he who can give it meaning.

Procedure

In order to investigate the behavioral meaning of deception, a study was made wherein two groups of subjects were recruited, adults and students. The sample consisted of ninety-six adults from the Minneapolis-St. Paul areas as well as one-hundred-seventy-eight students from the School of Business Administration at the University of Minnesota.

Each group of subjects was exposed to a series of television commercials which were unofficially rated for deception by four judges at the FTC. The ads submitted for review consisted of pairs of ads, each pair for the same product, with each member of the pair being a different ad. Pairs of ads utilized as stimulus material had to meet two criteria: (1) that at least three out of four judges agreed that one member of a pair was deceptive while the other ad was non-deceptive; and (2) at least three out of four judges agreed that no additional information was needed from the company to make a final judgment. Five pairs of ads met these criteria; in addition two control ads (one deceptive and one non-deceptive) in which at least three out of the four judges agreed were also selected.

Two reels of films were prepared (seven ads per reel); each reel contained either the non-deceptive or the deceptive version of the five experimental pairs of ads, as well as the non-deceptive and deceptive control ads. The two control ads were placed first, followed by the five experimental ads. The sequencing of ads for each reel was deceptive, non-deceptive, deceptive, non-deceptive, etc.

Subjects were randomly divided into two groups with one group seeing one reel of film while the other group saw the other reel. Both groups saw the same control ads which were utilized to determine if the groups perception of deception was comparable.

After viewing an ad, the subjects were asked to complete a checklist instrument. This instrument contained 49 dichotomous items reflecting the feelings of a large sample of consumers to a variety of advertisements (Bauer & Greysen, 1968). The checklist items were selected to be content free (apply to any ad), and included positive, negative, neutral, and ambivalent statements about the ad. A check made by a statement indicated that it reflected a subject's feelings about an ad. A large number of statements were included so as to insure that a respondent could reflect any particular feeling he had about an ad.

A series of phi inter-correlation matrices was developed for the checklist items within the adult and student samples and also within pairs of ads for each product category. This procedure yielded five inter-correlation matrices for each sample, or a total of ten. Each correlation matrix was tested to determine if it was significantly different from an identity matrix before the factor analyses were conducted (Weiss, 1970). All matrices were significant at the .001 level.

The principle factor method of factor analysis was utilized to analyze the ten phi inter-correlation matrices. The results of the factor analysis were then rotated, using the verimax orthogonal technique in order to gain better interpretability of the factors.

Results

An average of eight factors emerged in each of the ten factor analyses with approximately 88 percent of the common variance being explained by each analysis.

The major factors that appeared across student and adult samples were (1) enjoyment, (2) deception, (3) monotony, (4) information, (5) dislike, and (6) brand loyalty. The enjoyment factor dealt primarily with statements that indicated the presentation of the ad was well received. The deception dimension contained statements such as, "The ad was misleading," or that the claims made were unbelievable. The monotony factor contained statements concerning the sameness of the ad as compared with other ads. The informative dimension had statements such as, "It told me how to use the product," and "The ad told about new product uses," that loaded highly on it. The dislike factor included a variety of statements, such as "it was boring," or "it was annoying," and "it was poorly written or performed." Finally, the brand loyalty dimension had two basic sets of statements loading on it: one set indicated that the brand was previously used and liked; while the other set contained positive statements about the product, such as "it was informative."

The "pure" factors were not always found, however. The basic factors tended to mix or combine with one another yielding new factors, such as an

enjoyment-information factor, and an intelligence-monotony factor.

The deception factor was no different. Instead of clearly defined deception factors, a deception-intelligence, a deception-annoyance, and a deception-offensive factor emerged. The data indicated that the deception construct was not always unidimensional in nature as the legal meaning of deception implied.

Implications

The data from the study pose an important question. If a commercial insults a person's intelligence or annoys him, does this automatically lead an individual to perceive an ad as deceptive? In other words, does an individual view an ad as insulting to the intelligence and then make a deception judgment without ever attempting to make an independent evaluation of ad content? If this is so, it presents some serious problems for any consumer data that might be introduced into a deception case. The results of such data could not be meaningful unless the intelligence, annoyance, or offensiveness of an ad were controlled.

Footnotes

1. This paper is based on the author's dissertation, "The Perception of Deception in Television Advertising, An Exploratory Investigation." For more complete information on this paper, please write the author at the Department of Advertising, University of Illinois.
2. 382 U.S. 46 (1965).
3. Development in Law--Deceptive Advertising. Harvard Law Review, 1967, 80, p. 1076.
4. Developments in Law--Deceptive Advertising, p. 1077.
5. CCH Trade Regulation Reports. Washington, DC.: Commerce Clearing House, Inc., 1971, para. 7530. The complete definition of deception as well as all court citations are provided.
6. FTC v. Hires Turner Glass Co., CCH ¶ 7530 (CA-3; 1935) 1932-1939 TRADE CASES ¶ 55, 083, 11 F.2d 362 (FTC Dkt. 1985).
7. FTC v. Balme, CCH ¶ 7530 (CA-2; 1928) 23 F.2d 615 (FTC Dkt. 764).
8. John C. Winston, Co. v. FTC, CCH ¶ 7530 (CCA-3; 1925) 3 F.2d 961 (FTC Dkt. 1060).
9. Aronberg t. a. Positive Products Co. v. FTC, CCH ¶ 7530 (CCA-7; 1942) 1940-1943 TRADE CASES ¶ 56, 324, 182 F.2d 165 (FTC Dkt. 3856).
10. Exposition Press, Inc. v. FTC, CCH ¶ 7530 (CA-2; 1961) 1961 TRADE CASES ¶ 70, 146, 295 F.2d 869 (FTC Dkt. 7489).
11. John C. Winston Co. v. FTC, CCH ¶ 7530 (CCA-3; 1925) 3 F.2d 961 (FTC Dkt. 1060)
12. Dkt. 8538, CCH ¶ 7530 (CA-9; 1964) 1964 TRADE CASES ¶ 71, 278, 337 F.2d 751.

13. Ford Motor Co. v. FTC, CCH ¶ 7530 (CCA-6; 1941) 1940-1943 TRADE CASES ¶ 56, 132, 120 F.2d 175 (FTC Dkt. 3005).
14. Dkt. 8601, CCH ¶ 7530.
15. Gelb v. FTC, 144 F.2d 580, 583 (2d Cir. 1944).
16. FTC v. Winsted Hosiery Co., CCH ¶ 7530 (U.S. Sup. Ct. 1922) 258 U.S. 483 (FTC Dkt. 214)

References

- Bauer, R. & Greyser, S. Advertising in America: The Consumer View. Boston: Division of Research, Harvard University, 1968, pp. 197-235.
- Gellhorn, E. Proof of Consumer Deception Before the Federal Trade Commission. Kansas Law Review, 1969, 17.
- Kinter, E. W. A Primer on the Law of Deceptive Practices. New York: The Macmillan Company, 1971, pp. 31-34.
- Weiss, D. Factor Analysis and Counseling Research. Journal of Counseling Psychology, 1970, 17, pp. 477-485.

CONSUMER RESPONSES TO UNIT PRICING, OPEN DATING, AND NUTRIENT LABELING

Monroe Peter Friedman¹
Eastern Michigan University

This paper reviews recent research relating to three new informational displays which have been proposed to assist consumers with their shopping decisions. These innovations in labeling practices are unit pricing, open dating, and nutrient labeling. The literature reviews will be followed by a discussion of the relevance of the research findings to public policy decisions relating to the adoption of these proposed aids.

Unit Pricing

A number of terms (unit pricing, price-per-measure, dual pricing, and value pricing) have recently been employed to describe the practice of providing price information to consumers by such standard measures as the ounce, pound, pint, and quart. Unit pricing, as the practice is now widely called, is not a new development, having been used by supermarkets for many years to display price information for such store-packaged, variable-size products as meats and cheeses. What is new however is the effort to extend this practice to many thousands of pre-packaged retail products.

Unit-pricing advocates point to a host of packaging practices (packaging-to-price, cents-off labels, slack fill, and fractional quantity units) which they feel impair the ability of the consumer to compare prices effectively. They further contend that the tremendous increase in items carried by supermarkets (from about 1,000 in 1930 to 8,000 in 1970) has exacerbated the price-comparison problems of the consumer to an unacceptable level. These advocates believe that unit pricing will alleviate these problems by shifting the major share of the burden of price comparisons from the consumer to the retailer.

Research on unit pricing has been primarily concerned with three questions:

1. Is there a need for unit-price information?
2. How do consumers react to unit-pricing programs?
3. What are the costs of such programs?

Let us examine these questions in turn.

The Need for Unit-price Information

Studies addressed to this question have examined the ability of consumers to make price comparisons without the aid of unit-price labels. Four such studies have been reported (Friedman, 1966; Friedman, in press; Gatewood and Perloff, in press; and Houston, 1972), each of which instructed subjects to perform hypothetical price-comparison problems in a laboratory or field setting. All four studies revealed significant departures from perfect performance. In three of the four (Friedman, in press; Gatewood and Perloff, in press; and Houston, 1972), a second experimental condition was introduced to examine the value of unit-price information in solving the problems; all three found a substantial improvement in performance under the unit-price condition.

Taken together, these several studies reveal that without additional information, consumers are often unable to make correct price comparisons. The findings suggest that the use of unit-price labels may substantially reduce purchase errors for economy-minded consumers at a savings in time (Gatewood and Perloff, in press) and money (Friedman, in press; and Houston, 1972).

Consumer Reaction to Unit-pricing Programs

Monroe and LaPlaca (1972) have recently reviewed studies of consumer reaction to unit-pricing programs which were introduced in five food chains. Except for the common finding that reported usage of unit-price labels was correlated positively with education, the results of the five studies yield a most inconsistent picture. For example, two of the chains detected no shift in aggregate warehouse movements to products with lower unit prices, while a third found increased sales for the usually lower-priced private label brands. This result is in conflict with consumer survey responses in four of the chains which revealed changes in brands and/or sizes purchased by users of unit-price information. Reported usage of unit-price labels also varied considerably for the respondents in the five chains, with a high of 65 percent of the respondents in one chain and a low of 7 percent in another.

A number of methodological differences among the several studies may well be responsible for the conflicting findings. Likely candidates here are the unit of measure used to express unit-price information, the manner in which the information was displayed, the orientation program used to acquaint store customers with unit pricing, the proportion of store products which were unit-priced, and the time interval which elapsed before customer surveys were conducted.

Costs of Unit-pricing Programs

Several attempts have been made to measure the retailer costs of installing and operating unit-pricing programs. After reviewing six such studies, Monroe and LaPlaca (1972) conclude that the costs remain relatively constant per store, regardless of store size and sales volume. Thus small independent stores with low sales volume must expend a relatively high percentage of their annual sales to support a unit-pricing program. The estimates calculated in one study (McCullough and Padberg, 1971) of the direct costs of unit pricing as a percentage of total sales ranged from 4.15 percent for low-volume stores (annual sales under \$100,000) in a small distribution network (20 stores), to .095 percent for high-volume stores (annual sales over \$2,000,000) in a large distribution network (90 stores). It is important to note that these are direct costs, and do not consider the procedural efficiencies which may well result for retailers as a byproduct of instituting unit-pricing programs. In particular, three chains which have adopted programs report that they provide "tighter inventory control, better space management, and fewer price-marking errors." (Monroe and LaPlaca, 1972, p. 22).

Open Dating

Open dating refers to the practice of printing dates on packaged food products for the purpose of informing consumers about the freshness of the products. Advocates of the practice contend that consumers have a right to know whether the packaged food which they purchase is fresh. They further contend that open dating provides the necessary information to make this determination and thus should be adopted by the food industry to assist consumers with their shopping decisions. That the efforts of these advocates have received substantial public support is evident from the findings of a recent consumer survey study which was reported at the 1971 meeting of the National Association of Food Chains (Ringler and Berner, 1971). 89 percent of an enormous nationwide sample of 250,000 shoppers reported that they were in favor of easy-to-understand date codes to provide assurance of product freshness.

The food industry response to open-dating proposals has focused on a number of technical arguments which question the usefulness of dates as indicators of product freshness. These arguments are based, in large part, on the results of a recently completed two-volume study of food stability which was conducted by the Department of Food Science at Rutgers University (Food Stability Survey, 1971). The Rutgers' report presents extensive findings and recommendations as well as specific data for 18 consumer products. Perhaps its most significant general finding is that for most foods, temperature is a far more important determinant of product quality than elapsed time from the date of packaging. For example, the report finds that a frozen food product may deteriorate as much in two months at 10° F as it will in twelve months at 0° F. Thus freshness dates based simply on temporal considerations are not likely to be accurate unless temperature is carefully controlled at all stages of the food distribution cycle.

Advocates of open dating acknowledge the significance of possible deviations from proper temperature controls and handling procedures. They insist however that these problems can be overcome and point to the current widespread use of open dating by American food chains as evidence of its feasibility.

Among the many questions raised by the open-dating issue is the type of date information to be presented in open-dating programs. Some food processors have argued for a pack date since not only is it a known date, but it could be relatively easily provided at the time of packaging. Some retailers, on the other hand, have expressed a preference for a pull date, so that they would know the last acceptable date of sale. Moreover, some consumer advocates have suggested that a quality assurance date be adopted so that consumers would know the last date a product could be used at peak quality. Other consumer advocates have recommended an expiration date which would inform consumers of the last date on which safe usage of a product can be expected.

Of the more than 60 food chains which have recently adopted open dating, almost all have opted for pull dates. This choice recognizes the fact that proper temperature controls and handling procedures are more likely to be assured in the food processor-to-retailer stages of the food distribution cycle than in its later retailer-to-home-consumption stages. The choice also recognizes the value to the retailer of having simple last-day-of sale indicators to facilitate the removal of non-fresh products from store shelves and to help store clerks with stock rotation. Many retailers also see pull-date benefits for the consumer, believing, as they do, that this indicator of product freshness offers her the best combination of useable and reliable information.

Research on Open Dating

The economic costs of installing and operating an open-dating program have been examined by the U.S. Department of Agriculture and a study report is expected to be released in the near future. Other studies have looked at consumer awareness, understanding, and usage of open-dating information. Zehner (1971), in a survey study of 1,800 Michigan consumers, found that 80 percent of the participants had heard about open dating even though, at the time of the study, only a few Michigan food chains were operating open-dating programs. Zehner also showed future-dated packages to her respondents and asked them what was meant by the dates. The largest proportion (35%) interpreted them to be expiration dates while the second largest (26%) thought they were pull dates. When asked what they would like open dates to represent, 27 percent expressed a preference for expiration dates, 22 percent for a guide to freshness, 19 percent for pack dates, and 12 percent for pull dates. It is especially interesting to note that the least preferred indicator (pull dates) is the one most widely used by food retailers.

In a second study on open-dating, Taylor (1971) interviewed 1,700 female shoppers at 18 stores of a Chicago food chain which had practiced open-dating for several months by printing pull dates on some 150 private brand items. Slightly more than half of the respondents reported that they were aware of the dating information. Awareness was higher in middle-income areas (57%) than in low- and high-income areas (each about 50%).

About half (429) of the shoppers who had indicated awareness of the pull dates agreed to be interviewed in greater depth. Almost two-thirds (63%) of these shoppers reported they had used the dating information at least once. However one wonders about the success of their efforts in light of the finding that only 20 percent of the 429 were able to correctly interpret the pull date as the last day a product can be sold. Finally, when asked to cite advantages and disadvantages of the open-dating program, about half of the user group (54%) reported that "product freshness" was the principal advantage and 80 percent reported no disadvantages.

A third investigation of open-dating is being conducted by the Consumer Research Institute and a final report is expected to be published shortly. The report will present findings on a series of three studies. The first of these studies compares consumer reactions to pull dates and pack dates, each of which was introduced in three or more stores of a national food chain. The second study uses national mail and telephone surveys to examine consumer reactions to food quality and freshness. The third study in the series will provide before-and-after data relating to new open-dating ordinances which have recently been passed in New York City and Dade County, Florida. An effort will be made to determine the effects of the legislation on consumer attitudes and usage patterns.

To sum up, the few behavioral studies which have been conducted on open-dating raise more questions than they answer. Apparently, open-dating information is highly desired by American consumers (indeed, according to the Ringler and Berner study, more so than a host of other supermarket practices such as discounting, nutritional advice, and unit pricing). Yet expiration dates, which are most preferred by consumers, according to the Zehner study and some preliminary findings of a Consumer Research Institute study (Hoofnagle and Stokes, 1971), would appear to be the most difficult form of open-dating information for the food industry to provide. Other questions concern the confusion relating to the meaning of pull dates by the respondents in Taylor's study and, in light of this confusion, the relatively few disadvantages of the program which were reported by its users. A final unresolved question is the likelihood of selective consumer purchasing of late-dated food products. If, as has been suggested, many consumers would sort through the various packages on sale in supermarkets in an effort to buy the very freshest, many perfectly good food products would spoil on the store shelves, with a resultant economic loss to the retailer and the consumer. Although the Rutgers study states that the U.S. and foreign experiences with open-dating has found few instances of such selective purchasing, the authors cite no evidence to support this statement.

Nutrient Labeling

Prior to the White House Conference on Food, Nutrition, and Health, which was held in Washington in December of 1969, only modest efforts had been made to evaluate nutrient labeling as an aid to purchasers of packaged foods. The Conference, which was called "a watershed in American social history" by its chairman Jean Mayer (1972), in a report to the scientific community, broke from the virtual inactivity of the past by recommending that the Food and Drug

Administration (FDA) consider the development of a system for identifying the nutritional qualities of packaged foods. The Conference further recommended that the food industry be encouraged to provide nutritional information on package labels to enable consumers to follow recommended dietary regimens.

These Conference recommendations reflect the increasing number of processed and formulated food products which have recently appeared in the marketplace, most of which are not easily classifiable into standard nutritional categories. As Mayer has stated:

...Consumers regard labels as their primary source of information about the contents, safety, and nutritional value of the food they buy. Given our existing nutrition education programs, it is not difficult for the average citizen to select simple meats, vegetables, and fruits, and produce a balanced diet. But what if the choice involves such new foods as frozen pizza or spinach souffle? How should these products be classified, and how should they be labeled? (1972, p. 240)

The FDA responded to the Conference recommendation on nutrient labeling by proposing criteria for information panels on food packages in the Federal Register of March 30, 1972. These criteria specify the nutrients to be listed on food packages, the units by which they are to be quantified, and the permissible locations on food packages for presenting this information. If the criteria are formally adopted by the FDA, all nutritional information which appears on food packages will be required to adhere to these regulations. In the ninety-day period for filing written comments relating to this proposal, over 2,000 were submitted. They are now being processed and incorporated into a revised document which is expected to be published later this year.

Let us look briefly at the contents of the proposal and the research findings from which they derived. Since almost without exception, these findings are from unpublished reports, our review will of necessity be largely limited to the summaries which have been presented in the Federal Register in support of the proposed regulation. The proposed regulation calls for the prominent display of the following information on all food packages which present nutritional information:

1. A definition of serving size
2. Calorie content per serving
3. Number of grams of protein, fat, and carbohydrates per serving, as well as the amount of protein per serving expressed as a percentage of the 65 gram Recommended Daily Allowance (RDA)
4. The amount per serving of seven vitamins and minerals expressed as percentages of their RDAs. The seven are vitamin A, vitamin C, thiamin, riboflavin, niacin, calcium, and iron.

Research on Nutrient Labeling

The research which led to the proposed regulation began with a mail survey by Call and Hayes (1970) of 793 members of the American Institute of Nutrition, which sought their professional opinions on nutrient labeling. The study found that 85 percent of these nutritionists favored more nutrient information on the labels of food packages, with a majority giving highest ratings to the disclosure of calories, protein, fat, vitamins A, C, and D, calcium, iron, and additives and preservatives. While no consensus emerged on presentation modes for communicating nutrient information (e.g., per serving, per package, or per unit of weight), a majority expressed the opinion that vitamin and mineral disclosures should be based on a standard such as the RDA.

In a second study on nutrient information which was published in a trade magazine (Chain Store Age, October 1970), it was found that small shifts in consumer purchase patterns in the direction of nutrient-labeled food had followed their introduction in the marketplace.

Using this information as well as the advice and counsel of various professional, business, and consumer groups, the FDA proceeded to carry out additional research on nutrient labeling. Two series of investigations were initiated, one by the Consumer Research Institute, and the other by David Call and Daniel Padberg of Cornell University.

The first in the series of Consumer Research Institute studies was undertaken to secure information on presentation modes and label formats, as well as an indication of consumer understanding and usage of nutrient label information. This study examined the reactions of 950 educated, middle-class households which shopped by catalog, to the inclusion of nutrient information in the catalog descriptions of various food products. The preliminary findings revealed the following:

1. Purchase patterns shifted toward brands with nutritional advantages.
2. Additional consumer reactions to the catalog labeling program included a more positive attitude toward nutrition as well as greater nutritional knowledge.
3. Three experimental formats for disclosing the amount of RDA for the listed nutrients (numerical percentage, adjective representation, and pictorial representation), yielded no differences in consumer understanding or usage.
4. Differences in consumer reaction were also not found for two modes of displaying information for individual nutrients which are not present in a product. In one mode the nutrient was listed with a zero value, while in the other it was not listed at all.

A second study in the Consumer Research Institute series further explored consumer reactions to the three experimental formats and two display modes which were examined in the first study. This second study consisted of large-scale mail surveys of a U.S. probability sample of 2,000 consumers as well as a second sample of 2,000 low-income consumers. A third phase of this study used personal interviews to survey a sample of 600 customers who had not completed high school.

The preliminary findings revealed a slight overall advantage for the numerical percentage format over the pictorial and adjective formats - an advantage which was more pronounced for the sample of high school dropouts. Also of significance was the high percentage of respondents who were able to perceive differences as well as make correct nutritional choices using any of the formats (80 percent for the national sample, 70 percent for the low-income sample, and 90 percent for the low-education sample). With regard to the two display modes, a small advantage was found for lists which included only those nutrients present in a product, over lists which included all nutrients. However no attempt was made to determine the possible educational benefits to consumers of the more complete display mode.

Looking next at the Call and Padberg studies, we find that this effort was undertaken to secure a better understanding of consumer use and knowledge of nutrient labeling as well as consumer interest in nutritional information. Also sought was an indication of possible nonuse benefits to consumers. A second study was initiated to evaluate various nutrient label formats and to determine how a good diet may result from using nutrient information.

The preliminary findings which have been reported from the first of these studies are sketchy, representing early results from nutrient labeling experiments initiated by two food chains. The nutrient labeling programs of three additional food chains are also being evaluated by Call and Padberg, but at this writing, no findings have been reported.

To briefly summarize these various findings, it would appear that professional nutritionists agree on the need for more nutrient information of food packages as well as the specific nutrients to be listed. In addition, preliminary findings of three studies of various presentation modes and formats for nutrient information suggest that a variety of modes and formats can be effectively used and understood by consumers in general, and by disadvantaged consumers in particular, with a slight advantage appearing for numerical percentages of RDA over other forms of disclosure.

Policy Implications of Research on Consumer Aids

Although many significant insights have resulted from the research findings reported herein, their overall contribution to the policy-making process may be rather modest. This is so not because of specific deficiencies in the individual research studies, but rather because of the inherent limitations of the empirical research process as a contributor to the resolution of policy issues.

The questions for which answers are sought by policy makers concern the outcomes which are likely to follow the implementation of various policy proposals. Ideally, these answers would be expressed as a series of quantitative functions, each of which would indicate how much of a particular cost or benefit is likely to result at various points in time, from the date of adoption of a policy proposal to some fairly distant point in the future. This statement of ultimate criteria explicitly recognizes that 1) costs and benefits are multi-dimensional in nature, and 2) the long-term effects of new policies may differ markedly from their short-term effects.

With consumer aids, these two considerations are particularly important. The first emphasizes the necessity of going beyond simple frequency counts of aid users to a consideration of such additional criteria as consumer savings in time and money, nutritional improvements in diets, and both user and nonuser satisfaction with consumer aids. The second consideration suggests that since consumer aids are new and relatively unfamiliar to many shoppers, it may be several years before a true picture of stable usage and benefit patterns is likely to emerge. Indeed, when one considers the increasing attention being devoted to consumer economics courses in the public schools, it may be appropriate to speak of a generation gap with respect to public understanding of consumer aids which may take many years to close.

In light of the uncertainties inherent in these, as well as other developments, it seems clear that the future-oriented informational needs of the policy maker will remain unsatisfied. The methodological crystal ball of the social scientist, while not opaque, is still too cloudy to provide reliable predictions of future usage and benefit patterns for consumer aids.

Given the very real limitations in our forecasting powers, it is well to ask what role can be played by the social scientist in guiding policy decisions on consumer aid proposals. Two possibilities suggest themselves. The first consists of conducting longitudinal studies of shopper reactions to consumer aids, in an effort to establish at least a few critical data reference points from

which cost and benefit curves could be constructed and extrapolated into future time periods. Needless to say, there are many problems with this approach, not the least of which is the substantial waiting time before useful information is likely to result. A series of longitudinal studies conducted over a short period, say one to two years, is unlikely to provide a reliable basis for projecting consumer reaction over a longer period; and the myriad, real-world political pressures on policy makers often do not permit them to wait out a longer study period before formulating and acting upon public positions on proposed consumer aids.

A second and perhaps more promising approach to the problem of providing future-oriented information to policy makers draws a distinction between the actual and potential future value of information aids. Advocates of this approach believe that the social scientist can contribute little which bears directly on the former, but a great deal which sheds light on the latter. Thus they suggest that we do not attempt the near-impossible task of predicting the outcomes of policy decisions relating to consumer aids, but devote our efforts instead to determining the potential value of various proposed aids for the consumer.

But how, one may ask, do we define "potential value," and perhaps more importantly, how do we measure it for various consumer aids? A useful beginning may be made by viewing a consumer aid as having potential value if it presents useful information to the consumer in a highly communicable form. The determination of usefulness could be made by surveying various specialists, such as nutritionists and consumer economists, or consumers at large, although in some instances the latter group may lack the technical expertise to respond in a meaningful manner. Communicability, on the other hand, is more appropriately tested in a laboratory or field setting to assure that actual consumers are able to understand and effectively use the information provided by consumer aids.

Although many of the research studies reviewed in this paper have examined various aspects of information usefulness and communicability for consumer aids, they have typically failed to build upon one another to provide a meaningful picture of the potential value of the consumer aids under investigation. In large part this is due to the fact it is not the concepts of say, unit pricing or open dating, which have been researched in these many studies, but various particularities of these concepts. And with few exceptions, the particularizations selected for study have been determined by the exigencies of practical circumstance rather than the logic of a relevant theoretical model. As a result, it is difficult, if not impossible, to fit together what may be pieces from several different jigsaw puzzles into a unified whole.

The one exception to this pattern is the research on nutrient labeling. Due to the efforts of a single government agency, a programmatic approach to researching the potential value of this consumer aid has been undertaken which systematically explores the effects of various design parameters on different segments of the consumer population. While research of this sort will not furnish any ultimate answers to the policy maker, it may well serve the important function of helping him to decide which consumer aid proposals warrant serious consideration and which do not.

Footnotes

1. The author is professor of psychology and director of the Center for the Study of Contemporary Issues at Eastern Michigan University.
2. What may be operating here is something akin to what would seem to be the typical voter's attitude toward regulations requiring public disclosure of campaign contributions by political candidates. Few voters are likely to take the time to make sense of these disclosures but they may nevertheless favor such regulations knowing that the possibility of this information being used by even a small segment of the electorate may well serve to dissuade politicians from accepting unduly large contributions from single sources. If the reader will excuse a bad pun, the disclosure of information, even infrequently used information, may be favored by the public to assure that packaged foods can be safely bought and that political candidates cannot!

References

- Call, D. L. & Hayes, M. G. Reactions of nutritionists to nutrient labeling of foods. The American Journal of Clinical Nutrition, 23, 1970, 1347-1352.
- Food Stability Survey, Volumes I & II, Rutgers University, The State University of New Jersey, Department of Food Science, 1971.
- Friedman, M. P. Consumer confusion in the selection of supermarket products. Journal of Applied Psychology, 1966, 50, 529-534.
- Friedman, M. P. Consumer price comparisons of retail products: the role of packaging and pricing practices and the implications for consumer legislation. Journal of Applied Psychology, in press.
- Gatewood, R. D. & Perloff, R., An experimental investigation of three methods of providing weight and price information to consumers. Journal of Applied Psychology, in press.
- Hoofnagle, W. & Stokes, R. Consumer reactions to food quality in Conference Proceedings: Food Stability and Open Dating, October 21-22, 1971, Rutgers University, New Brunswick, N.J. 24-32.
- Houston, M. The effect of unit-pricing on choices of brand and size in economic shopping. Journal of Marketing, 36, 1972, 51-54.
- Mayer, J. Toward a national nutrition policy, Science, 176, 1972, 237-241.
- McCullough, T. D. & Padberg, D. I. Unit pricing in supermarkets: alternatives, costs and consumer reaction. Search Agriculture, 1, 1971, 1-25.
- Monroe, B. & LaPlaca, P. J. What are the benefits of unit pricing? Journal of Marketing, 36, 1972, 16-22.
- Ringler, M. & Berner, G. A. Consumer Attitudes Toward the Food Industry, paper presented at the 38th Annual Meeting of the National Association of Food Chains.
- Zehner, M. Consumer study on open dating. Conference Proceedings: Food Stability and Open Dating, October 21-22, 1971, Rutgers University, New Brunswick, N.J., 33-39.

SOURCE EFFECTS, MESSAGE EFFECTS, AND GENERAL EFFECTS

IN COUNTERADVERTISING

H. Keith Hunt
The University of Iowa

In January of 1972, the Federal Trade Commission (FTC) in a brief to the Federal Communications Commission (FCC) supported the concept of counteradvertising as a suitable approach to remedying some of the present failings of advertising. Counteradvertising is advertising which presents an opposing viewpoint to an ad that has already been run which the counteradvertiser finds objectionable in some way. This suggestion of the appropriateness of counteradvertising is the most recent in the series of actions taken by the FTC to inhibit the incidence of and offset the effects of false, misleading, and/or deceptive advertising on consumers and on competition in general. As with the previous introduction of corrective advertising, counteradvertising is proposed with no research evidence to support it.

Little is known about the effects of counteradvertising, even as to whether it has a positive or negative effect on attitude toward the brand. The FTC made its suggestion with no supporting evidence of the probable effects of the practice. The need for research into the effects of counteradvertising led to the following experiment designed to (1) investigate the effects of counterads, in general, on attitudes, (2) investigate the effects of different sources of counterads (consumer organizations, competitors, government agencies, and corrective ads by the original advertiser), and (3) investigate the effects of different message formats (unsupported statements as compared to supported, research-based statements). Findings from the investigation will give an indication both of the overall efficacy of counteradvertising and of which sources and message content choices are most effective.

The experiment exposed all subjects to the same advertisement and then introduced the various counterad treatments in the following 2 x 4 factorial design. Two message forms were used for the counterads: a supported statement and an unsupported statement. Four sources were used: a consumer organization, a competitor, a government agency, and the original advertiser. The original advertiser issued a "required corrective statement" which stated that the original ad was deceptive. The other sources offered the counter message under their name alone. Also, a control group was exposed to the allegedly deceptive ad only, providing a base for determining how much change the counterads caused. The experimental design is as follows:

		SOURCE				Control- No Counterad
		Consumer Organization	Competitor	Government Agency	Corrective Ad	
MESSAGE CONTENT	Supported					
	Unsupported					

Fig. 1.--The experimental design

It was expected, based on previous research on the effects of corrective advertising (Hunt, 1972), that the supported message would cause a significantly greater decrease in favorableness of attitude than would the unsupported message. Also, it was expected that the noncommercial sources would be more effective than commercial sources because of greater perceived truthfulness, and greater perceived sincerity.

Method

Subjects

A total of 270 subjects participated in the experiment and completed usable questionnaires. These were all unpaid students (sophomores, juniors, and seniors) in the Introduction to Marketing and the Computer Applications in Business courses at The University of Iowa.

Experimental Materials

The experimental materials consisted of several advertisements which were print ads of magazine quality, printed specially for this study.

The Allegedly Deceptive Ad. The allegedly deceptive ad used for the study was the original ad used to introduce Chevron's additive, F-310. This particular ad includes, according to the FTC, both misleading statements and deceptive presentation. The allegedly deceptive ad announces the new additive, F-310, and states that the additive reduces dirty exhaust emissions and thus improves gasoline mileage and helps toward cleaner air. Two pictures show balloons attached to cars' exhaust systems. The balloon attached to the car using Chevron with F-310 is clear. The other balloon is blackish from dirty exhaust emissions. This is offered as proof that Chevron with F-310 cleans the engine and reduces pollutant emissions (after being used for 2000 miles). Previous research into corrective advertising based on this ad found the ad to be easily understandable and meaningful to subjects (Hunt, 1972). Since Chevron products are not marketed in the geographic area where this study was run, the ad and counteradvertisements based on it would not affect the market for Chevron products. The openness of Chevron in allowing its controversial ad to be used for experimentation is appreciated.

The Chevron F-310 ad was used because it is an actual example of an ad counteradvertising would have been used to offset. Several groups were actively opposed to the ad before the FTC finally stepped in to require the corrective ad. The corrective ad requirement is currently in the hearing stage. It is possible that, except for the charge of false presentation, the rest of the claims made in the Chevron ad will not be disproved. Given the charges against the company's advertising and the controversial nature of the ad, Chevron is to be commended for its openness and willingness to allow the ad to be used in experimental studies.

The Counterads. Two forms of counterad were used, one form being an unsupported statement and the other being a supported, research-based statement. The unsupported statement has a headline reading "F-310 CLAIMS BY CHEVRON ARE FALSE AND DECEPTIVE ACCORDING TO (source)". Then, in the center portion of the page in large print is the following general statement: "TEST RESULTS SHOW THAT CHEVRON WITH F-310 DOES NOT CLEAN ENGINES AND DOES NOT REDUCE POLLUTION FROM UNBURNED HYDROCARBONS AND CARBON MONOXIDE EMISSIONS." At the bottom of the page is, "A PUBLIC SERVICE ADVERTISEMENT TO CONSUMERS FROM

(source)". Since no evidence is given to support the counterad statement and since the message was the same except for the source, any difference in effect would be source effect.

The supported research-based statement carried the same headline, "F-310 CLAIMS BY CHEVRON ARE FALSE AND DECEPTIVE ACCORDING TO (source)" but was followed by supposed research findings by independent research firms showing that Chevron with F-310 had no effect on reducing polluting exhaust emissions or cleaning the engine. These research findings were made up for this experiment and were not real. Four independent studies were referred to, each presenting evidence that Chevron with F-310 did not do what it claimed to do in its advertising. The closing statement was "THE EVIDENCE SHOWS THAT CHEVRON WITH F-310 DOES NOT REDUCE POLLUTION," followed by the same bottom line "A PUBLIC SERVICE ADVERTISEMENT TO CONSUMERS FROM (source)." Again, since the supported statements were the same except for the source, any differences in attitudes across sources would be source effects.

Since each source gave both a supported and an unsupported form of counterad, a comparison of results for each source would indicate the relative effectiveness of the supported as compared to the unsupported message format. Since the unsupported counterad had no evidence to support it, it was accepted primarily on the credibility of the source. The supported counterad had the same source effect plus the strength of the research findings from the four independent research firms. Thus, the supported counterad should be a stronger attack than the unsupported counterad, resulting in lower levels of attitude toward the deceptively advertised brand.

Measurement

Favorableness of attitude toward the brand was measured on the following 29 point scale. The subject expressed a degree of agreement or disagreement with a particular statement by circling the one dot in the line of dots which best represented his attitude concerning the statement.

I LIKE CHEVRON WITH F-310.

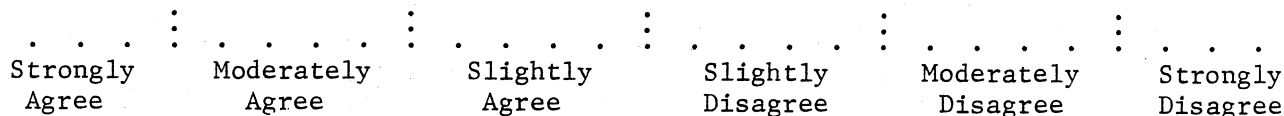


Fig. 2.--Measurement scale and basic affect statement.

In each case, the higher the score on the measurement scale, the stronger the agreement with the statement. The attitude statements for which measures were obtained are:

1. I like Chevron with F-310.
2. I think claims made by Chevron with F-310 were truthful.
3. I think Chevron is an expert source of information on gasoline and gasoline additives.

4. I think Chevron is sincere in its advertising about Chevron with F-310.
5. I think the claims made by (source) were truthful.
6. I think (source) is an expert source of information on gasoline and gasoline additives.
7. I think (source) is sincere in its advertising about Chevron with F-310.

General Procedure

The experimenter was introduced to the class as a professor who was doing a study of what people thought about different kinds of advertisements. The professor explained that the study would take about 15 minutes and that it involved carefully reading one or two ads and answering a short questionnaire about them.

The materials were in a 9 x 12 envelope which was passed out in a manner to assure that each cell was equally represented in each class group. The advertisements were stapled together in one packet with a cover and spacer sheets. The measurement scales were in another. Subjects were asked to take out the ad packet and turn to the first ad, reading it carefully until told to turn to the second ad. This took five minutes. Then the subjects turned to the second ad and read it for five minutes. Control subjects had no second ad (counterad) and were asked to wait for the next materials. Then all the ad packets were put back in the envelope and the measurement questionnaire taken out and completed. Students were told to put their names on their questionnaires. The questionnaires were put back in the envelope and the envelopes collected. Then the professor thanked the participants and left the room.

Statistical Procedure

Analysis of variance was used to test for significance of main effects. Tests of significance between cells in the experimental design were computed using Tukey's HSD test (Kirk, 1968). In the discussion section which follows, "not significant" means not significant at the .05 level of significance. "Significant" will have the level of significance indicated.

Results

The control group data indicate that exposure to the Chevron ad resulted in a slightly favorable attitude toward the brand. However, counteradvertising in all but one case led to a substantially and significantly less favorable attitude toward the brand. The only exception is the competitor's unsupported attack which is discussed later. The findings are as follows:

		SOURCE				
		Consumer Organization	Competitor	Government Agency	Corrective Ad	Control- No Counterad
MESSAGE CONTENT	Supported	9.82	9.96	9.19	10.44	17.23
	Unsupported	10.29	15.19	9.26	10.80	

Numbers above are cell means.

Within cells means square error = 35.021

Cell size = 27

Tukey's HSD values:

HSD Down .01 level = 4.15

.05 level = 3.16

HSD Across .01 level = 5.10

.05 level = 4.13

(If the difference between two cell means is greater than the HSD value, that difference is significant at the level of significance for that HSD value.)

Fig. 3.--Counteradvertising effects on affect toward Chevron with F-310

A presentation of the findings in graphic form is presented on the following page.

In answer to the first research question about the general effects of counteradvertising, counterads do substantially reduce favorableness of attitude toward the brand. In seven of the eight cells the attitude level was significantly ($p < .01$) less favorable after a counterad exposure than before. Only in cell 2,2 was the cell mean not significantly different from the pre-treatment, control group mean.

To answer the second research question about source effects, different sources of counterads did not make any difference except for the specific case of competitor-unsupported ads. While the source main effect was significant at the .05 level, this significance was due to the extreme variation in cell 2,2. The means were almost identical in the other seven cells.

To answer the third research question about message effects, supported and unsupported statements have the same effect except, again, in the case of cell 2,2, the competitor-unsupported ad, which was not significantly different from the control mean. While the message content main effect was significant at the .10 level, this significance was due to the extreme variation in cell 2,2,. As is seen in Figure 4, the means for each source, except for cell 2,2, were almost identical.

Except for one of the eight treatments, neither source nor message content differences made any difference in liking for the brand--each combination resulted in the same slightly to moderately unfavorable level of attitude toward the brand.

Discussion

Source Effect

The similarity of source effects is surprising. It was expected that the consumer organization and government agency, the noncommercial sources, would effect the greatest decrease in attitude, that the corrective ad would effect less of a decrease and that the competitor would effect the least decrease.

It seemed reasonable that the noncommercial sources would be perceived as having less personal interest in the matter and more of a public service orientation, and would also be perceived as more sincere and more truthful than the commercial sources. Only in the dimension of expertness were the commercial sources expected to have a favorable advantage. Based on these expectations the noncommercial sources were expected to effect a greater decrease in favorableness of attitude than the commercial sources. Yet the four resulted in essentially the same level of attitude.

A review of the measures of truthfulness, expertness, and sincerity, revealed that all the predictions were borne out. Noncommercial sources were perceived as more sincere and more truthful. After exposure to a counterad, non-commercial sources were even seen as slightly more truthful. But the end result was

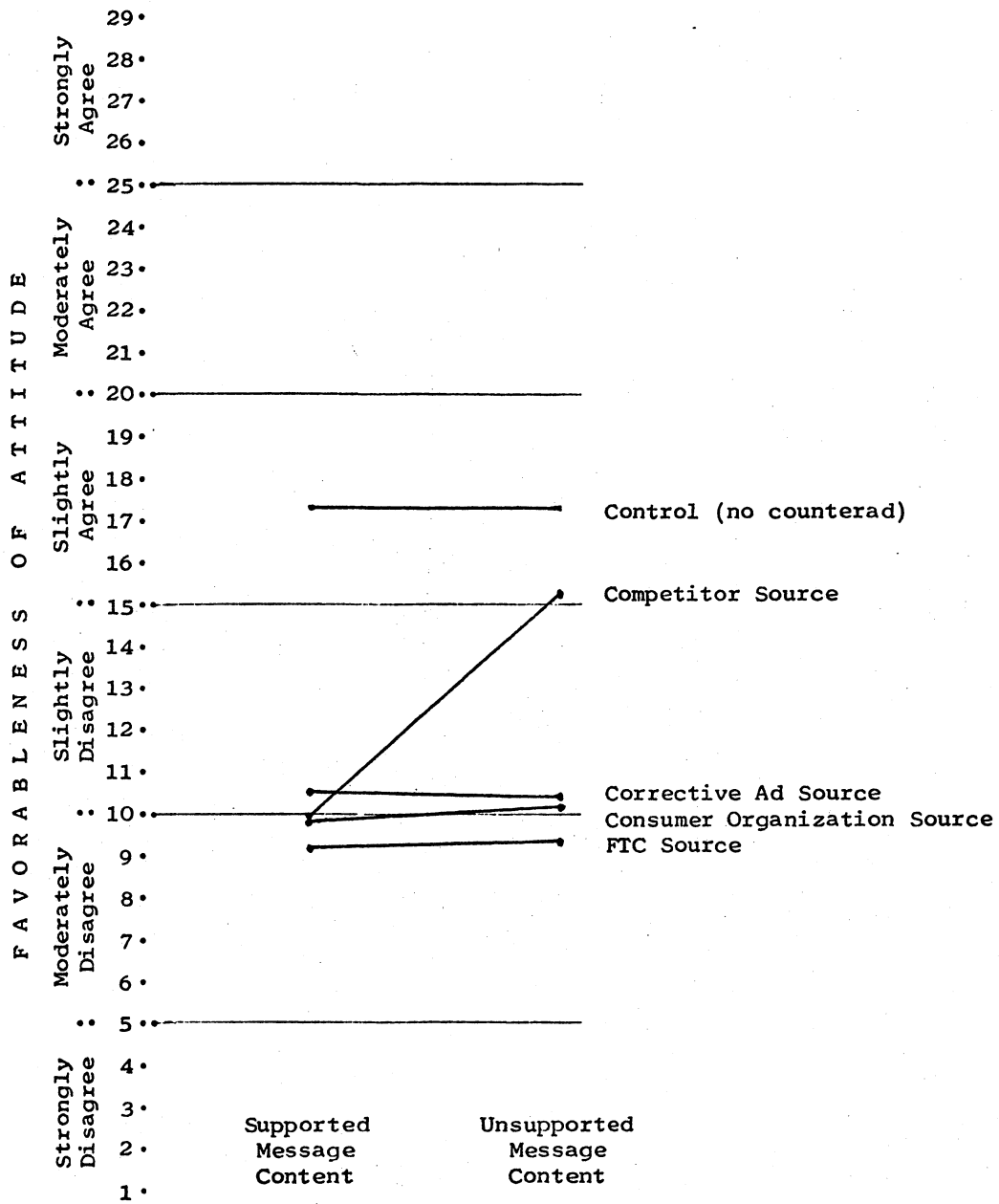


Fig. 4.--Effect of counteradvertising on affect toward Chevron with F-310.

nearly identical attitude levels for all sources. Further investigation is needed into which dimension or causes were strongest and what interactions occurred to yield this unexpected outcome.

Effects of Exposure to Counteradvertising on Perceived Truthfulness, Expertness, and Sincerity

It is interesting to note the changes in perceived truthfulness, expertness, and sincerity after exposure to the counterad. Perceived truthfulness of the sources was about the same before and after exposure to the counterad except for the corrective ad situation. Being required to admit in one's own ad that previous ad claims had been false caused a substantial drop in the perceived truthfulness of the corrective advertiser. After exposure to the counterad, perceived expertness rose for the noncommercial sources and dropped for the commercial sources.

Perceived sincerity dropped substantially for all sources of counterads, except the competitor, who wasn't perceived as very sincere in the first place. The noncommercial sources and the brand itself each dropped about equally in the sincerity rating. Evidently sources pay a price of decreased perceived sincerity when they engage in counteradvertising efforts. While expected for the brand, this is an unexpected finding for the noncommercial sources.

Previous research on the effects of corrective advertising showed perceived sincerity of the FTC to be unaffected by requiring a company to run corrective advertising (Hunt, 1972). Overall attitude toward the FTC actually increased after requiring the corrective ads. The decreased sincerity of both noncommercial sources is worth further investigation, but stands as a warning for the time being that counteradvertising may have deleterious effects for the counter-advertiser as well as the original advertiser.

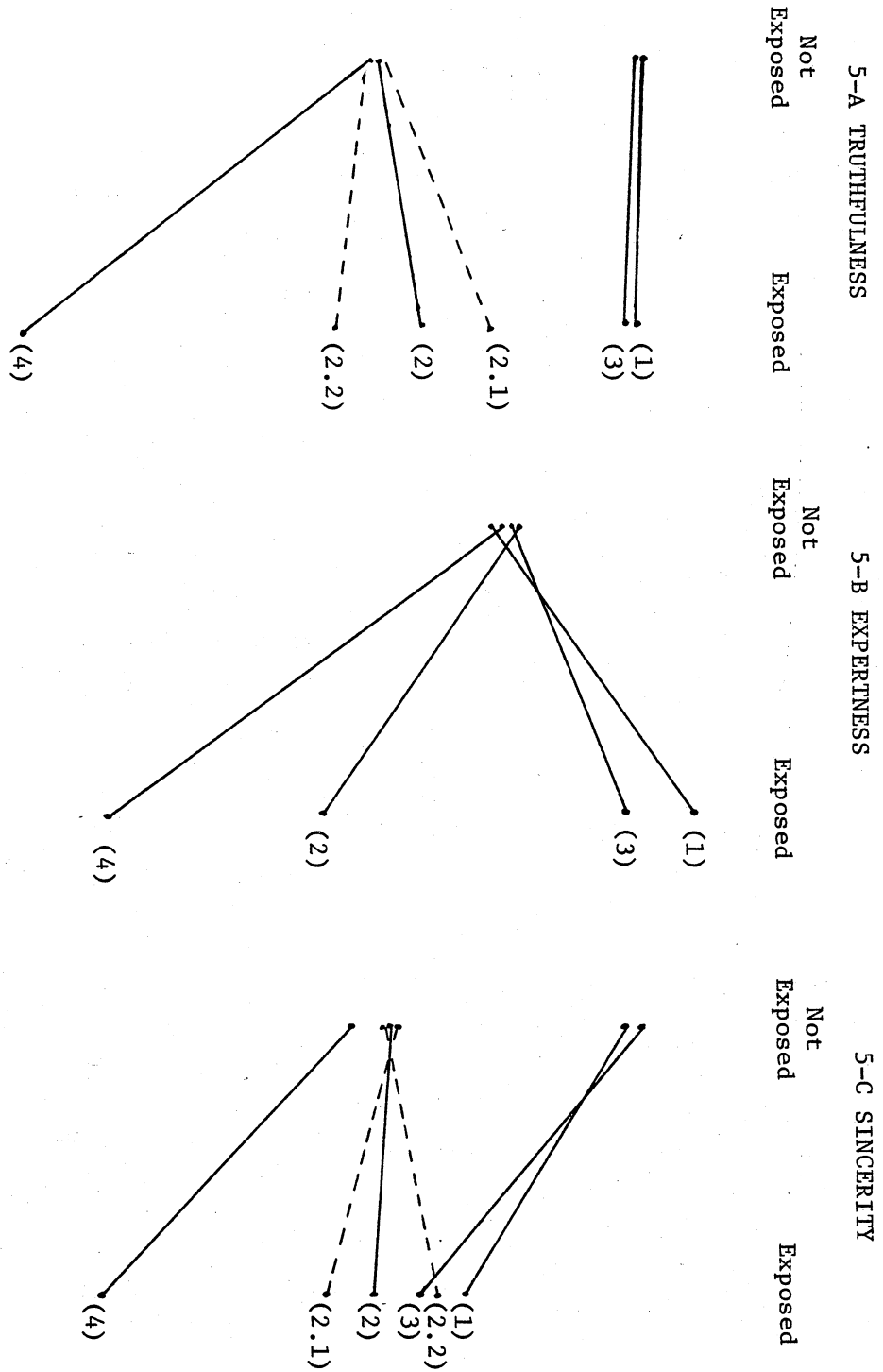
Message Effects

Previous research on corrective advertising established that the supported attack was much more effective than the unsupported attack in reducing favorableness of attitude. In both studies the supported statement has been a pointed, precise attack on the falseness of the Chevron advertising claims. Also, in both studies the unsupported statement has merely stated that previous advertising by Chevron was false and misleading but no evidence or other specifics were mentioned. In the corrective ad study overall attitude was very little different after the unsupported statement attack than when no attack occurred. But the supported attack resulted in a substantial decrease in attitude. In the current study a much lower initial favorableness occurs, with equal drops in attitude from both formats, but not as much of a drop as the supported previously obtained. That the experimental situation was somehow different is likely, but subjects were similar, and ad formats were very similar. Nevertheless, the similar effects of both message formats is unexpected. It doesn't seem to make any difference which message is used.

The Competitor-Unsupported Case

The substantial differences in attitude level evoked by the competitor's counterads, with supported being much more effective than unsupported, is of interest. The measure of economic interest shows the competitor was considered very differently than the noncommercial sources--being perceived as interested in counteradvertising for his own advantage and not especially for the public interest. The result is that the competitor is a noncredible source not effective in reducing favorableness of attitude very much. Yet the supported attack

Strongly Disagree 5
 Moderately Disagree 10
 Slightly Disagree 15
 Slightly Agree 20
 Moderately Agree 25
 Strongly Agree



- Key:
- (1) Consumer Organization
 - (2) Competitor
 - (2.1) Competitor, Supported Statement
 - (2.2) Competitor, Unsupported Statement
 - (3) Government Agency
 - (4) Corrective Ad by Chevron

Fig. 5 -- Effects on perceived truthfulness, expertness, and sincerity of exposure to a counterad.

Note: For ease of presentation, the average of the two cell means for each source (supported and unsupported) is used. When a source's cell means differ substantially, both means are shown.

Table 1
Effect of Exposure to a Counterad on Perceived Truthfulness

	Not Exposed to Counterad				Exposed to Counterad			
	Consumer Organization	Competitor	Government Agency	Chevron Corrective Ad	Consumer Organization	Competitor	Government Agency	Chevron Corrective Ad
supported	21.41	16.52	21.56	16.82	23.11	18.89	22.30	8.96
supported	22.93	16.15	22.70	15.59	20.89	15.44	21.33	7.78
IX	22.2	16.3	22.1	16.2	22.0	17.2	21.8	8.4

Effect of Exposure to a Counterad on Perceived Expertness

Table 2

		Not Exposed to Counterad				Exposed to Counterad			
		Consumer Organization	Competitor	Government Agency	Corrective Ad	Consumer Organization	Competitor	Government Agency	Corrective Ad
	supported	18.33	18.56	19.93	18.15	23.78	14.93	22.63	9.67
	unsupported	19.19	20.11	18.22	19.82	22.56	15.07	20.89	10.59
	X	18.8	19.3	19.1	19.0	23.2	15.0	21.8	10.1

Table 3
Effect of Exposure to a Counterad on Perceived Sincerity

	Not Exposed to Counterad				Exposed to Counterad			
	Consumer Organization	Competitor	Government Agency	Corrective Ad	Consumer Organization	Competitor	Government Agency	Corrective Ad
supported	21.19	16.59	21.33	16.44	18.19	15.04	17.30	10.11
unsupported	22.48	16.33	22.78	15.04	18.04	17.56	17.07	9.89
IM	21.8	16.5	22.1	15.7	18.1	16.3	17.2	10.0

was very effective. This strong effect is attributed to the research findings and their sources which are explicitly discussed in the ad. The four research studies, each by a group independent of the advertiser, were evidently so highly believable that they overwhelmed the competitor source effect. It is supposed that if the research had been done by the competitor rather than by independent research firms, the statements would have been less effective in reducing attitude. In the unsupported attack condition, no external source existed and the acceptance of the message rested solely on the credibility of the competitor, and the result was very little effect.

This previous point raises the question of whether the exceptionally strong supported statement with its four external sources not only swamped the competitor source effect but all four source effects. That all four sources had such similar attitude levels certainly supports such a contention. However, for the corrective source and both noncommercial sources, the unsupported attack based only on the credibility of the counteradvertiser source resulted in attitude level equally as low as the supported attack. It seems to be more a case of agreement between sources, counterad and research, than swamping, except in the case of the competitor-supported, in which the competitor source, being weak, was evidently swamped by the research sources mentioned in the statement.

Conclusion

It was expected (1) that counteradvertising would cause a decrease in favorableness of attitude, (2) that supported counter statements would be more effective than unsupported counter statements, and (3) that different sources would have different effects. Only the first expectation was borne out.

Several points should be kept in mind when interpreting and generalizing the present results. The study works with attitudes rather than purchase behavior, with some imitation rather than real ads, with lengthy forced single intensive exposures rather than several short happenstance exposures, with an unfamiliar product rather than a product currently available and being purchased, and with one product rather than several product categories. Also, students are far from being a cross section of gasoline purchasers. However the corrective advertising study found student responses to be very similar to results obtained from a cross section of the buying public. Finally, as presently conceptualized, counteradvertising would occur only in broadcast media--radio and television. Print ads were used in the study because of familiarity with the Chevron F-310 campaign and to allow for possible comparisons with the corrective advertising studies. The print ads were also more orderly to administer. Any studies eliminating one or more of these limitations would be a contribution to understanding the effects of counteradvertising. Hopefully, the government agencies involved in counteradvertising will some day feel a commitment to backing similar studies with their expertise and funds, and, as a result, basing their proposals on research findings as much as possible.

References

- Hunt, H. Keith. Deception, Inoculation, Attack: Implications for Inoculation Theory, Public Policy, and Advertising Strategy. Unpublished Ph.D. dissertation, Northwestern University, June, 1972.
- Kirk, Roger E. Experimental Design: Procedures for the Behavioral Sciences. Belmont, California: Brooks/Cole Publishing Company, 1968.

THE COMPONENTS OF PERCEIVED RISK

Jacob Jacoby
Department of Psychological Sciences
Purdue University
and Leon B. Kaplan
Advertising Research Department
E. I. DuPont

Since Bauer's seminal statement (1960), a considerable number of researchers have utilized the construct of perceived risk to investigate various aspects of consumer behavior. Unfortunately, this construct has more often than not been defined in different ways by different investigators. As a result, the risk literature contains numerous instances where findings of one study either conflict with or cannot be directly compared with those of another.

This state of affairs has been partially remedied by Roselius (1971), who incorporated numerous risk reduction strategies (e.g., attend to product rating services; engage in word-of-mouth conversations; solicit information from sales personnel; read relevant mass-media ads; etc.) into one study and established their relative order of importance. Yet, while we have begun to learn the manner in which commonly used risk reduction strategies are related to each other, the interrelationships existing among the various kinds of perceived risk (e.g., functional risk, psychological risk, financial risk, etc.) are still unknown. To the extent that one investigator manipulates one kind of risk and another utilizes a second type, it becomes difficult to meaningfully compare and assess the significance of these independent studies.

Accordingly, the purposes of this investigation were: (a) to identify the various types of risk referred to in the literature; (b) to determine their interrelationships; and (c) to determine their individual and collective relationship to overall perceived risk.

Procedure

Identifying the Varieties of Perceived Risk

The consumer behavior literature on perceived risk was reviewed and the different varieties of perceived risk employed, whether as components of overall perceived risk or as operational definitions of it, were cataloged. Differences in terminology were disregarded when it was clear that the terms were functionally equivalent. In conjunction, a series of hypothetical purchase situations were conceptually developed and examined in terms of the types of risk potentially operative in each situation. Five types of perceived risk emerged from these procedures to subsume the types of risk found in the literature and generated by the hypothetical purchasing situation. These were: financial, performance, physical, psychological, and social risk. On a conceptual level, these five dimensions can be considered functionally independent so that as one risk variety increases, the other risk varieties can either increase, decrease, or remain unaffected. Thus, while psychological and social risk are usually fused and treated as one (i.e., psycho-social risk), the former should

probably be reserved for situations regarding how the individual perceives himself while the latter used to refer to the consumer's perception of how others will react to his purchase. As an example, if a 20-year old uses Geritol and does not tell anyone about it, it may affect the way he thinks of himself but would not necessarily affect the way others think of him.

It is interesting to note that these five risk varieties can be inferred from Bauer's original work (cf. 1960, p. 390). Moreover, considering overall perceived risk (OPR) as consisting of several independent varieties of risk suggests situations in which the consumer will engage in risk "trade-off" behavior. Buying a high-whitening (and possible highly abrasive) toothpaste may decrease social risk while increasing physical risk. Similarly, the housewife purchasing an expensive cut of meat for a dinner party she is giving is decreasing performance and social risks while increasing financial risk.

Assessing the Interrelationships Among the Varieties of Perceived Risk

Operationalizing the varieties of perceived risk. Table 1 contains the operational definitions of the five varieties of risk and of OPR. These

TABLE 1

Operational Definitions of the Varieties of Perceived Risk

<u>Type of Perceived Risk</u>	<u>Operational Definition</u>	<u>Anchor Points</u>
1. Financial Risk	What are the chances that you stand to lose money if you try an unfamiliar brand of _____ (either because it won't work at all, or because it costs more than it should to keep it in good shape)?	1=low chance of losing money; 9=high chance of losing money
2. Performance Risk	What is the likelihood that there will be something wrong with an unfamiliar brand of _____ or that it will not work properly?	1=low functional risk; 9=high functional risk
3. Physical Risk	What are the chances that an unfamiliar brand of _____ may not be safe; i.e., may be (or become) harmful or injurious to your health?	1=very safe; 9=very unsafe
4. Psychological Risk	What are the chances that an unfamiliar brand of _____ will not fit in well with your self-image or self-concept (i.e., the way you think about yourself)?	1=low psychological risk; 9=high psychological risk
5. Social Risk	What are the chances that an unfamiliar brand of _____ will affect the way others think of you?	1=low social risk; 9=high social risk
6. Overall Perceived Risk	On the whole, considering all sorts of factors combined, about how risky would you say it was to buy an unfamiliar brand of _____?	1=not risky at all; 9=extremely risky

definitions were constructed to be as basic and as uniform as possible, and were judged to be clear enough to permit trained subjects to have similar ideas as to what each of the types of risk meant. Cunningham's (1967, pp. 104-105) notion of general versus specific risk levels were also considered. The definitions are general, though a specific object (an unfamiliar brand) is the focus of the rating. Substitution of a brand name without altering the structure of the operational definition is easily accomplished.

Subjects. The subjects, 148 upper-classmen at Purdue University enrolled in Consumer Psychology during the spring 1970 semester, were not naive, having just completed a two-hour section on the topic of perceived risk, with each of the risk varieties defined as in Table 1.

Instrument. A questionnaire was developed to measure both the amount of each specific type of perceived risk as well as the overall perceived risk these subjects associated with 12 different consumer products. Products were intuitively selected to cover a substantial portion of the overall perceived risk continuum. An assortment of health, recreational, and hygienic products, varying along an expensive-inexpensive dimension and including products bought for self and for others, highly visible and low profile items, intimate and non-intimate products were utilized. The only restriction was that the products be appropriate for both sexes. The number of products was limited to 12 to reduce respondent fatigue and to minimize the introduction of error.

Each page of the questionnaire assessed one component of perceived risk separately for each of the 12 products. The top of each page contained the definition of each component.

Results

Table 2 contains the means and standard deviations for the risk ratings within each product category. The products are ordered in terms of mean OPR value. The ranking tends to be consistent with what would be expected if the products were ordered simply by price alone. In addition, a certain degree of construct validity was established in that the order of risk components was the same within meaningful clusters of products. For example, the rank order of the risk components for the three products which were items of apparel (i.e., suit, overcoat, and dress shoes) was identical and went from social risk, through psychological, financial, and performance risk to physical risk. Similarly, the rank order of components for three of the drug products (i.e., toothpaste, vitamins, and aspirins) was the same and went from physical risk through performance, financial, and social risk to psychological risk. It would seem that similar types of products have similar risk component hierarchies.

The product categories ranked by their mean values on each of the risk components are displayed in Table 3. Examination of the financial risk column indicates that the mean values go through a gradual diminution until dress shoes, then there is a drop in mean value of 1.59 (5.14 to 3.55) for deodorants, and then the means continue their modest decrementation. This pattern has some interesting implications for the relationships between actual price and perceived financial risk. While there seems to be a monotonic relationship between actual price and perceived financial risk, it appears to be curvilinear. At the extreme upper end of the price continuum

TABLE 2

Means and Standard Deviations for Perceived Risk
Ratings Within Each Product Category*

FOREIGN SPORTS CAR			LIFE INSURANCE			23" TABLE COLOR TV		
	\bar{X}	SD		\bar{X}	SD		\bar{X}	SD
Overall	7.25	1.53	Overall	7.01	1.84	Overall	6.26	1.87
Financial	7.76	1.57	Financial	7.17	1.81	Performance	6.32	2.05
Social	7.61	1.80	Performance	6.72	1.99	Financial	6.30	1.92
Psychological	7.18	2.16	Psychological	4.91	2.53	Social	6.11	2.07
Performance	7.05	1.73	Social	4.78	2.41	Psychological	5.07	2.24
Physical	5.24	2.32	Physical	4.45	2.82	Physical	3.47	2.20
SUIT			WINTER COAT			PAIR OF DRESS SHOES		
	\bar{X}	SD		\bar{X}	SD		\bar{X}	SD
Overall	5.91	1.74	Overall	5.39	1.81	Overall	5.18	1.78
Social	7.31	1.88	Social	6.64	1.98	Social	6.29	2.18
Psychological	6.91	1.92	Psychological	6.25	2.34	Psychological	6.06	2.32
Financial	6.41	1.82	Financial	5.53	2.12	Financial	5.14	1.93
Performance	5.84	1.96	Performance	4.97	2.13	Performance	4.93	2.09
Physical	2.08	1.54	Physical	2.78	1.82	Physical	3.45	2.06
DEODORANTS			RAZOR BLADES			TOOTHPASTE		
	\bar{X}	SD		\bar{X}	SD		\bar{X}	SD
Overall	4.02	2.07	Overall	3.39	1.83	Overall	3.00	1.81
Performance	4.68	2.31	Performance	3.99	2.21	Physical	4.05	2.19
Physical	4.07	2.17	Physical	3.80	2.16	Performance	3.24	1.98
Social	3.70	2.42	Financial	3.09	1.84	Financial	2.47	1.55
Financial	3.55	2.22	Psychological	2.29	1.73	Social	2.28	1.54
Psychological	3.25	2.18	Social	2.14	1.62	Psychological	2.03	1.37
VITAMINS			ASPIRIN			DECK OF PLAYING CARDS		
	\bar{X}	SD		\bar{X}	SD		\bar{X}	SD
Overall	2.87	1.71	Overall	2.50	1.72	Overall	1.61	1.10
Physical	4.25	2.29	Physical	3.91	2.43	Social	2.22	1.79
Performance	3.01	1.97	Performance	2.76	2.31	Financial	1.87	1.42
Financial	2.64	1.84	Financial	1.72	1.05	Psychological	1.86	1.41
Social	1.97	1.44	Social	1.60	1.22	Performance	1.83	1.40
Psychological	1.41	1.44	Psychological	1.47	1.27	Physical	1.27	0.73

*Measured on a 1-9 scale, where 9 = high perceived risk.

TABLE 3

Ranking of Products by Mean Risk Values for Each Type of Perceived Risk

	\bar{X}	SD		\bar{X}	SD
OVERALL:			PERFORMANCE:		
1. Sports Cars	7.25	1.53	1. Sports Cars	7.05	1.73
2. Life Insurance	7.01	1.84	2. Life Insurance	6.72	1.99
3. Color TV	6.26	1.87	3. Color TV	6.32	2.05
4. Suits	5.91	1.74	4. Suits	5.84	1.96
5. Winter Coats	5.39	1.81	5. Winter Coats	4.97	2.13
6. Dress Shoes	5.18	1.78	6. Dress Shoes	4.93	2.09
7. Deodorants	4.02	2.07	7. Deodorants	4.68	2.31
8. Razor Blades	3.39	1.83	8. Razor Blades	3.99	2.21
9. Toothpaste	3.00	1.81	9. Toothpaste	3.24	1.98
10. Vitamins	2.87	1.71	10. Vitamins	3.01	1.97
11. Aspirins	2.50	1.72	11. Aspirins	2.76	2.31
12. Playing Cards	1.61	1.10	12. Playing Cards	1.83	1.40
Mean:	4.53			4.61	
FINANCIAL:			SOCIAL:		
1. Sports Cars	7.76	1.57	1. Sports Cars	7.61	1.80
2. Life Insurance	7.17	1.81	2. Suits	7.31	1.88
3. Suits	6.41	1.82	3. Winter Coats	6.64	1.98
4. Color TV	6.30	1.92	4. Dress Shoes	6.29	2.18
5. Winter Coats	5.53	2.12	5. Color TV	6.11	2.07
6. Dress Shoes	5.14	1.93	6. Life Insurance	4.78	2.41
7. Deodorants	3.55	2.22	7. Deodorants	3.70	2.42
8. Razor Blades	3.09	1.84	8. Toothpaste	2.28	1.54
9. Vitamins	2.64	1.84	9. Playing Cards	2.22	1.79
10. Toothpaste	2.47	1.55	10. Razor Blades	2.14	1.62
11. Playing Cards	1.87	1.42	11. Vitamins	1.97	1.44
12. Aspirins	1.72	1.05	12. Aspirins	1.60	1.22
Mean:	4.47			4.39	
PSYCHOLOGICAL:			PHYSICAL:		
1. Sports Cars	7.18	2.16	1. Sports Cars	5.24	2.32
2. Suits	6.91	1.92	2. Life Insurance	4.45	2.82
3. Winter Coats	6.25	2.34	3. Vitamins	4.25	1.97
4. Dress Shoes	6.06	2.32	4. Deodorant	4.07	2.17
5. Color TV	5.07	2.24	5. Toothpaste	4.05	2.19
6. Life Insurance	4.91	2.53	6. Aspirins	3.91	2.43
7. Deodorants	3.25	2.18	7. Razor Blades	3.80	2.16
8. Razor Blades	2.29	1.73	8. Color TV	3.57	2.20
9. Toothpaste	2.03	1.37	9. Dress Shoes	3.45	2.06
10. Vitamins	1.91	1.44	10. Winter Coats	2.78	1.82
11. Playing Cards	1.86	1.41	11. Suits	2.08	1.54
12. Aspirins	1.47	1.27	12. Playing Cards	1.27	.07
Mean:	4.10			3.58	

TABLE 4

Ranking of Correlations Between Type of Risk and Overall Perceived Risk for Each Product

Foreign Sports Cars		Life Insurance		23" Color Tv		Suits		Winter Coats		Dress Shoes	
1. Performance (.490)	1. Financial (.607)	1. Performance (.598)	1. Performance (.563)	1. Performance (.596)	1. Social (.555)						
2. Financial (.451)	2. Performance (.470)	2. Financial (.482)	2. Social (.516)	2. Social (.561)	2. Performance (.525)						
3. Social (.411)	3. Psychological (.341)	3. Psychological (.479)	3. Psychological (.457)	3. Psychological (.537)	3. Psychological (.479)						
4. Psychological (.359)	4. Physical (.234)	4. Social (.354)	4. Financial (.408)	4. Financial (.467)	4. Financial (.451)						
5. Physical (.211)	5. Social (.218)	5. Physical (.142)	5. Physical (.142)	5. Physical (.275)	5. Physical (.142)						
Deodorants		Razor Blades		Toothpaste		Vitamins		Aspirins		Playing Cards	
1. Financial (.683)	1. Performance (.782)	1. Performance (.626)	1. Physical (.585)	1. Performance (.623)	1. Performance (.628)						
2. Performance (.675)	2. Financial (.581)	2. Physical (.612)	2. Financial (.564)	2. Physical (.557)	2. Financial (.506)						
3. Social (.657)	3. Psychological (.483)	3. Social (.542)	3. Performance (.564)	3. Social (.391)	3. Physical (.490)						
4. Psychological (.496)	4. Social (.453)	4. Psychological (.487)	4. Psychological (.490)	4. Psychological (.368)	4. Psychological (.475)						
5. Physical (.488)	5. Physical (.411)	5. Financial (.462)	5. Social (.399)	5. Financial (.312)	5. Social (.439)						

(foreign sports car and life insurance) actual dollar value is less important than is relative cost. The level of measurement seems to have shifted from interval to ordinal. The most meaningful price break (for the sample of college students used) should occur between shoes (\$15 - \$35) and the \$2 - \$3 items (deodorants, razor blades, etc.), and that appears to be what happened as evidenced by the discontinuity in Table 3.

An analogous situation would seem to exist for social risk. The most severe break in mean values occurs between color TV and life insurance (6.11 to 4.78). A color TV and the product rates above it on social risk (sports cars, suits, coats, and shoes) are all highly visible relative to life insurance and the remaining, lower-rated products.

Another observation which has implications for the validity of the measurement technique is that the three items of apparel cluster next to each other on every risk dimension except financial, and even there only one product (color TV) intrudes on the trio.

Table 4 contains the ranking of the correlations between OPR and its components. It is interesting to note that performance risk correlates highest with OPR in 8 of the 12 product categories, and second highest in another three categories. This would support the selection of performance risk as an approximation of OPR by many of the risk researchers. Financial risk has the highest correlation in two product categories (life insurance and deodorants), social risk in one (dress shoes), and physical risk in another (vitamins).

Table 5 contains the correlation between OPR and its hypothesized components across the 12 products examined in this study. As would be expected from Table 4, performance risk correlates highest with OPR (.654) and is followed closely by financial risk (.627).

TABLE 5

Correlations Between the Risk Components and
Overall Perceived Risk Across the 12 Product Categories

Performance	.654
Financial	.627
Social	.596
Psychological	.571
Physical	.552

Table 6 contains the summary tables of a stepwise multiple regression used to predict OPR from its component values. The multiple r 's range from a low of .6354 for foreign sports cars to a high of .8330 for razor blades. All are significant at $p < .0001$. This means that from 40% to nearly 70% of the variance associated with the OPR ratings can be accounted for by the five putative varieties of perceived risk outlined here. It also seems

TABLE 6

Summary Table for Predicting Overall Perceived Risk from Its Component Values

Step	Variable	Multiple r	Increase in r ²	Coefficient	Step	Variable	Multiple r	Increase in r ²	Coefficient
RAZOR BLADES									
1	Performance	.7820	.6116	.50639	1	Financial	.6830	.4665	.33926
2	Social	.8208	.0622	.20345	2	Social	.7916	.1602	.30867
3	Financial	.8302	.0156	.14944	3	Performance	.8282	.0592	.26295
4	Physical	.8328	.0042	.05953	4	Physical	.8303	.0035	.07246
5	Psychological (Constant)	.8333*	.0008	.04045 .14568)	5	Psychological (Constant)	.8316*	.0022	-.05876 .34273)
TOOTHPASTE									
1	Performance	.6306	.3976	.25586	1	Performance	.5955	.3546	.28590
2	Physical	.7075	.1030	.27944	2	Social	.7093	.1485	.27420
3	Social	.7482	.0592	.27660	3	Financial	.7266	.0249	.13348
4	Financial	.7677	.0296	.22986	4	Physical	.7390	.0181	.12877
5	Psychological (Constant)	.7681*	.0006	-.04232 -.08393)	5	Psychological (Constant)	.7489*	.0148	.12156 .28887)
ASPIRINS									
1	Performance	.6228	.3879	.32430	1	Performance	.6283	.3948	.27114
2	Physical	.7010	.1035	.21806	2	Physical	.6930	.0856	.39689
3	Social	.7448	.0633	.31430	3	Psychological	.7280	.0497	.13880
4	Psychological	.7470	.0033	.09805	4	Financial	.7373	.0136	.11743
5	Financial (Constant)	.7470*	.0001	-.01604 .12399)	5	Social (Constant)	.7416*	.0064	.06171 -.00063)
DEODORANTS									
WINTER COATS									
DECK OF PLAYING CARDS									

*p < .0001

TABLE 6 (continued)

Step	Variable	Multiple r	Increase in r ²	Coefficient	Step	Variable	Multiple r	Increase in r ²	Coefficient
VITAMINS									
1	Physical	.5691	.3239	.2777	1	Performance	.5625	.3164	.31647
2	Financial	.6990	.1646	.25800	2	Social	.6768	.1416	.29392
3	Psychological	.7296	.0438	.21519	3	Psychological	.6986	.0300	.15440
4	Performance	.7323	.0039	.08658	4	Financial	.7097	.0157	.13347
5	Social	.7348*	.0036	.08457	5	Physical	.7113*	.0022	.05480
	(Constant)			.18067)		(Constant)			-.12694)
PAIR OF DRESS SHOES									
1	Social	.5552	.3083	.28623	1	Performance	.5976	.3571	.36705
2	Performance	.6770	.1501	.23931	2	Financial	.6299	.0396	.21537
3	Financial	.6948	.0244	.16843	3	Psychological	.6445	.0187	.07891
4	Psychological	.7052	.0146	.11299	4	Physical	.6505	.0077	.07103
5	Physical	.7052*	.0000	.00477	5	Social	.6543*	.0049	.07817
	(Constant)			.63022)		(Constant)			1.45566
LIFE INSURANCE									
1	Financial	.6068	.3682	.47552	1	Performance	.4899	.2400	.25951
2	Psychological	.6351	.0352	.10735	2	Social	.5861	.1035	.22512
3	Performance	.6464	.0144	.13477	3	Financial	.6251	.0473	.21811
4	Physical	.6487	.0030	.03156	4	Psychological	.6345	.0119	.08482
5	Social	.6495*	.0010	.02785	5	Physical	.6355*	.0012	.02385
	(Constant)			1.89196)		(Constant)			1.27798)
23" TABLE COLOR TV									
FOREIGN SPORTS CARS									

*p < .0001

that as price rises, the r_m^2 's decrease. This suggests that some price-related aspect of OPR is not being accounted for by the five components and, as price rises, this variable becomes increasingly important in the estimation of OPR. This price-related aspect of OPR would seem to be independent of financial risk because financial risk enters the regression equation earlier for the last six products (mean value = 2.5) than for the first six products (mean value = 3.3).

Table 7 summarizes the results of trying to predict OPR from its components across all 12 product categories. The five components accounted for 61.6% of the variance in OPR ($p < .0001$). In stepwise regression that variable correlating most highly with the criterion enters the equation first. This obviously was performance risk. In the second step, the variable accounting for most of the residual variance is entered. Social risk was the second variable to enter the regression equation. In a like fashion, financial, physical, and psychological risk were entered.

TABLE 7

Summary Table for Predicting Overall Perceived Risk
from Its Component Values Across All 12 Product Categories

<u>Step</u>	<u>Variable</u>	<u>Multiple r</u>	<u>Increase in r^2</u>	<u>Coefficient</u>
1	Performance	.6536	.4272	.21259
2	Social	.7415	.1226	.43572
3	Financial	.7644	.0344	.23685
4	Physical	.7779	.0208	.13488
5	Psychological	.7848*	.0108	.12263
	(Constant			10.43503)

*p < .0001

Comparing Table 7 with Table 5, which contains the mean correlation between each component and OPR, it can be seen that social risk and financial risk have reversed positions, as did physical and psychological risk. Table 8, the intercorrelation matrix of the components, explains the discrepancy. The highest intercorrelation between variables (.655) occurs between performance and financial risk and, of the remaining variables, social risk has the lowest correlation with performance risk (.428). The second highest correlation between components is between social and psychological risk (.605). This probably explains why psychological risk entered the equation last. Interestingly, all the coefficients of Table 8 are significant at $p < .01$, implying that, at least as operationalized in this investigation, the components of perceived risk are not independent.

TABLE 8

Intercorrelation Matrix for the Five Types of Risk,
Calculated Across the 12 Product Categories

	Psychological	Performance	Physical	Social
Financial	.395	.655	.448	.441
Psychological		.511	.441	.605
Performance			.461	.428
Physical				.477

Discussion

Perhaps the most basic finding of this investigation is that overall perceived risk can be predicted fairly well from five putative components -- financial, performance, physical, psychological, and social risk. The resulting r^2 's ranged from 40% to 70% (all significant at $p < .0001$), with a median of $r^2 = .55$ across the 12 product categories. These results suggest that the varieties of perceived risk and overall perceived risk are related in a manner such as is depicted by the following formula:

$$\text{OPR} \approx f(\text{Uncertainty of Financial Risk X Consequences of Financial Risk;} \\ \text{Uncertainty of Performance Risk X Consequences of Performance} \\ \text{Risk;} \\ \text{Uncertainty of Physical Risk X Consequences of Physical Risk;} \\ \text{Uncertainty of Psychological Risk X Consequences of Psycho-} \\ \text{logical Risk;} \\ \text{Uncertainty of Social Risk X Consequences of Social Risk}); + \\ \text{error.}$$

In addition to its predictive validity, these analyses demonstrate construct validity. In examining the mean values for each component on a product-by-product basis (cf. Table 2), the similar ranking of the components for the three items of apparel (i.e., suit, winter coat, and dress shoes) was noted, as was the similarity for three drug products (i.e., toothpaste, vitamins, and aspirin). It would seem that similar types of products possess similar risk component hierarchies.

Finally, the basic significance of this investigation lies in the fact that it appears to represent the first comprehensive attempt to examine the interrelationship of the various types of perceived risk to each other as well as to overall perceived risk. Outlining these relationships should facilitate making meaningful and valid comparisons across investigations which employ different types of perceived risk.

Footnote

1. Since this study was conducted (Spring 1970), Roselius (1971, p. 58) has identified a sixth variety of risk: "Time loss: when some products fail, we waste time, convenience, and effort getting it adjusted, repaired, or replaced."

References

- Bauer, R. A. Consumer behavior as risk taking. In R. S. Hancock (Ed.), Dynamic marketing for a changing world. Chicago: American Marketing Association, 1960, 389-398.
- Cunningham, S. M. The major dimensions of perceived risk. In D. F. Cox (Ed.), Risk taking and information handling in consumer behavior. Cambridge, Mass.: Harvard University Press, 1967, 82-108.
- Roselius, T. Consumer rankings of risk reduction methods. Journal of Marketing, 1971, 35, 56-61.

PERCEIVED RISK: A MEASUREMENT METHODOLOGY AND PRELIMINARY FINDINGS

James R. Bettman
University of California, Los Angeles

Since Bauer (1960) introduced the concept of perceived risk, many research studies have explored the ramifications of this notion. However, research in this area has been hampered by the lack of a suitable measurement of perceived risk itself. The measures proposed to date have been arbitrary for the most part, and have often differed from one study to the next. This study proposes a measurement methodology for perceived risk and reports some initial research using this measure.

An initial measurement for perceived risk was proposed by Cunningham (1967), utilizing ordinal certainty and danger scales. Cunningham then arbitrarily combined these scales multiplicatively and developed risk categories. This final ordinal measure has been used by several other studies. Bettman (1969) proposed an ordinal measure of perceived risk based on paired comparisons. Both of the above measures suffer from having only ordinal properties; an interval-scaled measure would be more desirable, as it would allow a wider range of analytical tools to be employed in examining research results. Two different rating scale measures of risk that are assumed to be interval were proposed by Spence, Engel, and Blackwell (1970) and by Perry and Hamm (1969). Finally, some studies utilizing perceived risk notions have used very special purpose and arbitrary notions rather than attempting to develop a general measure of risk (For example, Cox and Rich, 1964 and Sheth and Venkatesan, 1968).

In view of this variety of measures and sometimes questionable methodology, a theoretical perspective for measurement is needed. First, a distinction should be made between two different types of risk, which have been confused in previous studies. Inherent risk is the latent risk a product class holds for a consumer, the innate degree of conflict the product class arouses in the consumer. Handled risk is the amount of conflict a product class engenders when the buyer chooses a brand from that product class in his usual buying situation. Thus handled risk includes the effects of information and risk reduction processes as they have acted on inherent risk. Cunningham's (1967) measure seems to deal with inherent risk. However, Cox and Rich (1964) and Spence, Engel, and Blackwell (1970) seem to be dealing with handled risk. The purpose of making the distinction between the two constructs is to allow for greater precision in measurement.

Measurement Methodology

The measures developed for inherent and handled risk are based on the notion of distinguishing the concepts by means of the choice situation utilized for measurement. In both cases an extended paired comparison method was used.

In each case, a brief description of the concept of risk, outlining the ideas of uncertainty, consequences, economic, and social risks, was given. Then, for inherent risk the S was asked, for pairs of product types, to rate which product type of the pair would be more risky to shop for in an imaginary store where all packages were labeled only with a product type and size, no brand labels or prices. Ss were told that this store had a full selection of all the usual brands for each product type; they were just to assume they could not tell which brand was which. After choosing which product type was more risky, Ss then rated how much more risky that type was on a scale ranging from 0 (equally

Table 1
Characteristics of Risk Distributions

Product type	Mean	Standard Deviation	Skewness	Kurtosis
Inherent Risk				
Paper Towels	-2.186	1.816	- .186	- .021
Dry Spaghetti	-1.432	2.387	1.146	1.631
Furniture Polish	- .594	2.782	.640	.056
Toothpaste	1.706	2.562	.309	- .526
Beer	.716	2.889	.025	.432
Instant Coffee	1.103	2.605	.039	- .232
Aspirin	.440	3.489	.264	- .569
Margarine	1.165	2.526	- .078	- .130
Fabric Softener	- .919	2.322	.503	- .095
Handled Risk				
Paper Towels	-1.178	1.966	.101	1.939
Dry Spaghetti	-1.121	2.039	1.173	2.645
Furniture Polish	.517	2.792	.333	- .237
Toothpaste	.304	2.490	.581	- .288
Beer	.706	3.028	.069	.574
Instant Coffee	.701	2.406	.582	.335
Aspirin	.259	2.910	.275	- .060
Margarine	- .125	2.183	.344	.031
Fabric Softener	- .061	2.427	.627	- .026

risky) to 9 (very much more risky). Handled risk was measured in essentially the same way, but by evoking a different choice situation. The S was asked to rate which product type in a pair was more risky to shop for in her own usual grocery store. Hence the effect of brand knowledge and information is included.

These measures were obtained from a convenience sample of 97 housewives, using nine product types: paper towels, dry spaghetti, furniture polish, toothpaste, beer, instant coffee, aspirin, margarine, and fabric softener. Scale values for inherent and handled risk for each of the nine product types were developed from the paired comparison data for each subject. The scaling method is quite simple. For the 36 product type pairs ($9 \cdot 8/2$) rated as outlined above, the rating for each ordered pair (i, j) is denoted by d_{ij} . The sign convention used is $d_{ij} > 0$ if i is more risky than j , and $d_{ij} < 0$ if j is more risky than i for the ordered pair (i, j). By convention $d_{ii} = 0$. Since only half of the off-diagonal pairs are rated, set $d_{ji} = -d_{ij}$. Then scale values are simply obtained by setting the scale value S_i for stimulus i to $S_i = \sum_j d_{ij} / 9$. Note that by using this procedure the S_i sum to zero. The mean has been removed. Hence we have relative measures for inherent and handled risk. The measures can range from -8 to $+8$ because of the rating scale used.

In addition to the proposed measures of inherent and handled risk, a modified version of Cunningham's (1967) certainty and danger questions was employed. Instead of using Cunningham's scale, a ten point rating scale was used for each component. For the certainty component, the scaling was then reversed from Cunningham's convention, so that in this study the scale ranged from 1 (very certain an untried brand will work as well as my present brand) to 10 (almost never certain it would work as well). Hence it is an uncertainty measure in this format. For the danger component, the convention utilized by Cunningham was used: 1 (no danger in trying a brand you never tried before) to 10 (a great deal of danger).

Results

Several sets of analyses were performed utilizing these measures. The methodology used for each analysis will be discussed in the appropriate section.

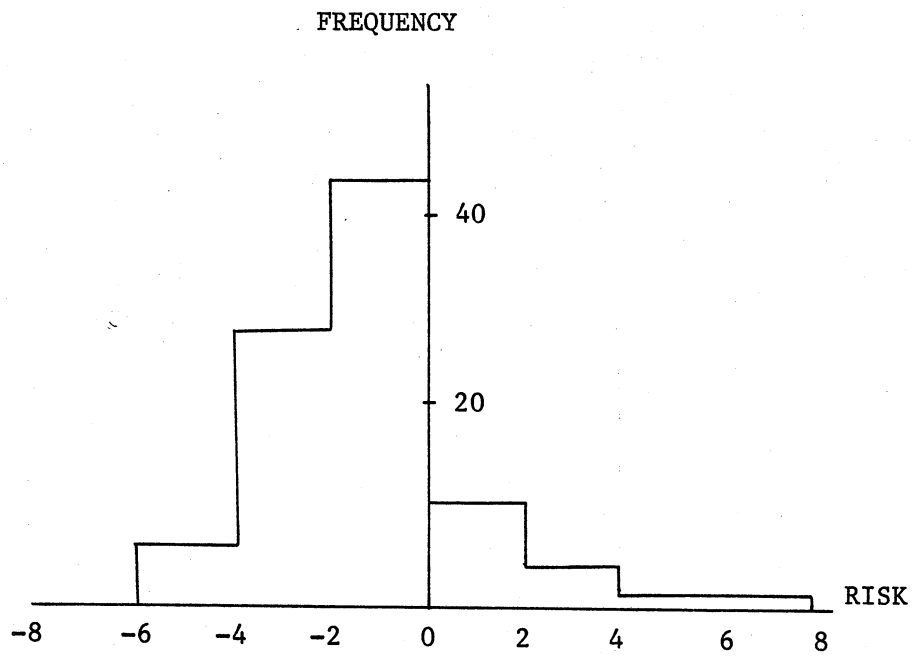
Perceived Risk over the Population of Subjects

Properties of the distributions of inherent and handled risk for each product type were computed, and are presented in Table 1. Since there are only 97 Ss, it is premature to authoritatively discuss these distributions. However, bearing this caveat in mind, note that toothpaste and margarine had the highest ratings for relative inherent risk, and beer and instant coffee had the highest ratings for handled risk. Aspirin and beer seem to have the highest variation in the risk scale values; this result seems intuitively plausible, since housewives presumably do not know much about brands of beer and perceptions of differences among brands of aspirin vary widely. Finally, note that there is a tendency for the curves to be skewed to the left. Typical histograms for handled risk are shown in Figure 1, and for inherent risk in Figure 2.

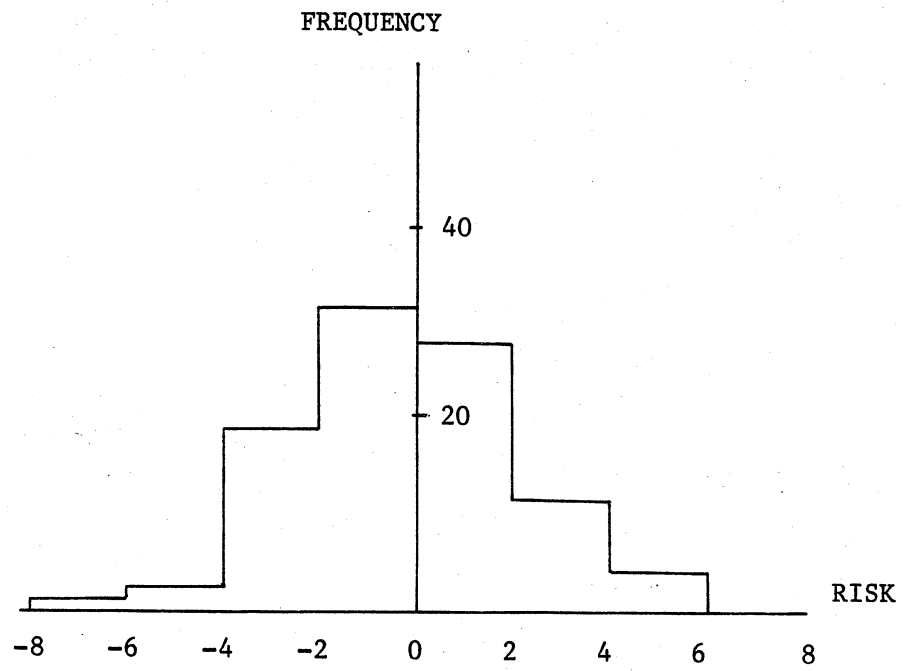
To examine the properties of the risk scales one step further, a cluster analysis using a modified K-Means algorithm was performed on the inherent risk scale values (McRae, 1970). Euclidian distance was used in the analysis. This analysis was also performed for handled risk, but the results were similar and

FIGURE 1

HANDLED RISK HISTOGRAMS



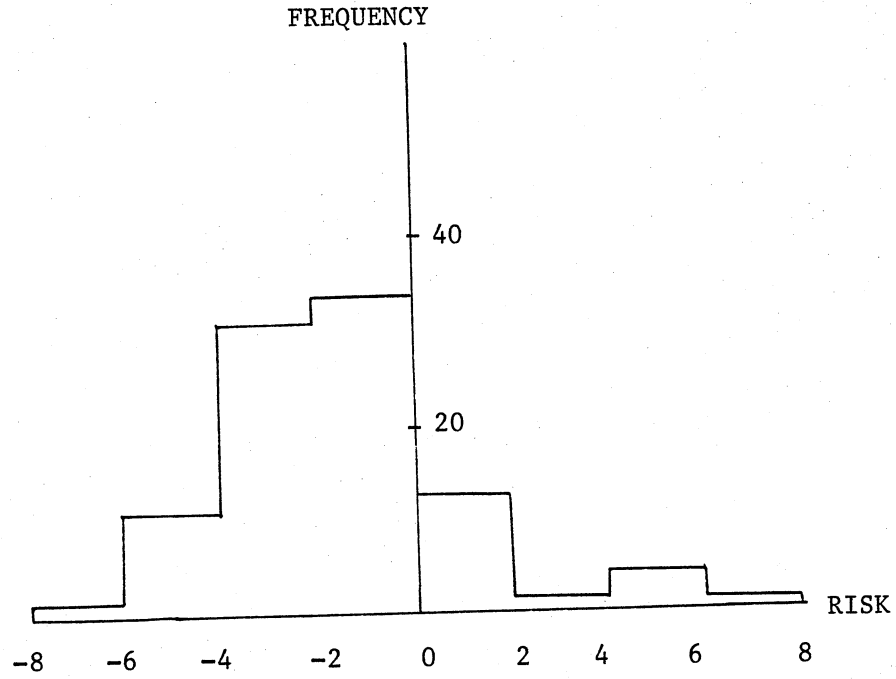
DRY SPAGHETTI



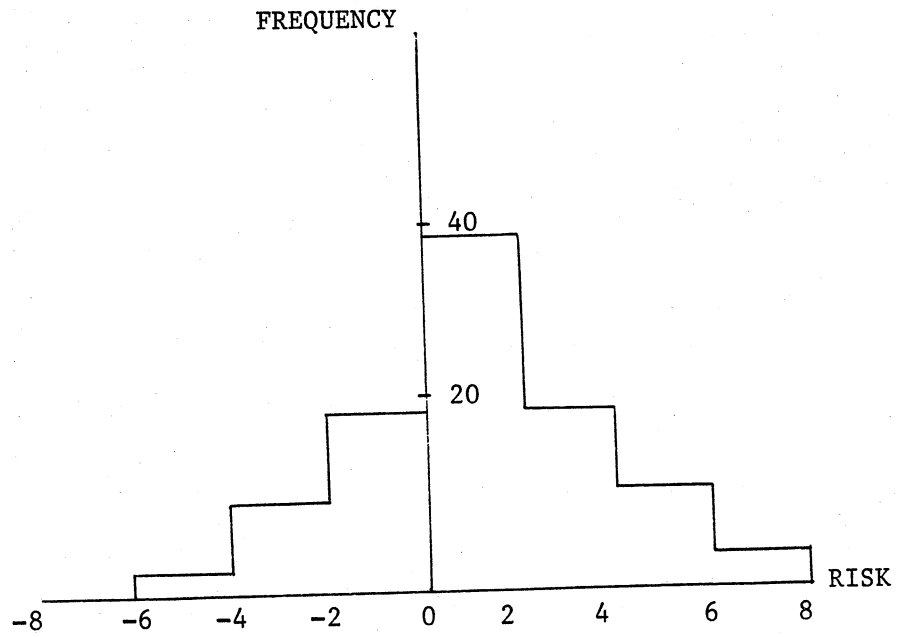
MARGARINE

FIGURE 2

INHERENT RISK HISTOGRAMS



DRY SPAGHETTI



MARGARINE

are not presented here. The cluster centroids for the six inherent risk clusters developed are given in Table 2. The clusters are quite interesting. For example, cluster 5 seems to associate risk with household chore products, whereas cluster 3 associates high risk with the socially oriented types of beer and instant coffee. More research needs to be done here to see if cluster membership can be predicted on the basis of S characteristics.

The Relationship between Inherent and Handled Risk

The second set of findings examines the relationship between the inherent and handled risk ratings. One hypothesis is that the inherent risk scales are more extreme than the handled risk scales, that they depart more from zero. This seems plausible on intuitive grounds; the less the information used in developing the rating, the more pronounced the tendencies to differentiate product types should be.

To test this hypothesis, the sums of the squared scale values for both handled and inherent risk were developed for each S. The hypothesis is then that the sum of squared values for inherent risk is greater than the sum for handled risk. By making several assumptions, it would be possible to show that these sums of squares approximate chi-squared variables, use the additive property of chi-square over Ss and form an F-ratio to test the equality of the sums of squares for inherent and handled risk. However, it was felt that the assumptions necessary were not justified. Hence, a more conservative alternative was taken. The sign test was used to test the hypothesis. If it rejects the null hypothesis that the sum of squares for inherent risk is equally likely to be either greater or less than the sum of squares for handled risk, we can be fairly confident in the result, since its power-efficiency is low. For the 97 Ss, 65 had higher sums of squares for inherent risk, 31 for handled risk, and there was one tie. The one-tailed probability of this result under the null hypothesis is .0005. Thus, we can reject the null hypothesis.

If the scale values for inherent risk are more extreme than those for handled risk, there are several ways this could occur. One possibility is that inherent risk values tend to expand uniformly away from zero compared to handled risk values. That is, negative and positive values of handled risk are associated with more negative and more positive values of inherent risk respectively. The ordering of product types on risk would stay the same, but the scales would expand outward from zero for inherent risk. One test of this hypothesis is to perform a simple regression analysis using inherent risk as the dependent variable and handled risk as the independent variable. If the regression weight were greater than one with an intercept of zero, this would support the hypothesis above. Two sets of nine product type regressions were run, one set for all 97 Ss and one set for just the 65 Ss for whom the sums of squared scale values were larger for inherent risk. In both cases, the fits of the regressions were good, with R^2 ranging from .177 to .662 for the 97 S runs and from .149 to .623 for the 65 S runs. However, in only one case, aspirin for the 65 S cases, was the regression weight greater than one. Hence, it seems clear that uniform expansion is not responsible for the more extreme values for inherent risk. As can be seen looking at the averages in Table 1, there is some switching of product type order occurring.

Relationships with Cunningham's Components

As noted above, Cunningham's certainty and danger measures were transformed to ten point rating scales for uncertainty and danger. After this transformation, both scales should be directly proportional to risk. Cunningham used an arbitrary

Table 2
Inherent Risk Cluster Centroids

Product Type	Clusters					
	1	2	3	4	5	6
Paper Towels	-2.278	-3.126	-2.833	-1.290	-1.244	-3.473
Dry Spaghetti	-3.185	- .172	-2.462	-1.206	-1.318	-1.100
Furniture Polish	.315	1.543	-2.612	-1.214	2.280	-3.300
Toothpaste	5.157	1.902	-1.740	.825	- .154	3.420
Beer	- .297	-3.124	6.647	1.257	.947	.377
Instant Coffee	2.499	-1.433	5.537	.979	.545	.860
Aspirin	-3.444	4.642	- .557	- .265	-1.974	4.082
Margarine	.879	.889	- .018	2.308	- .805	2.237
Fabric Softener	.352	-1.122	-1.963	-1.395	1.719	-3.100
N	12	9	6	26	21	23

multiplicative model for risk. However, parameters can be estimated for both additive and multiplicative models relating the scales derived from Cunningham's work to the inherent risk measure proposed in this study in an attempt to derive less arbitrary weightings for the two components. Inherent risk is used since it seems to be most related to the measures Cunningham employed. Thus, if Y represents inherent risk, X_1 uncertainty, and X_2 danger, the two models examined are the simple linear model:

$$(1) \quad Y = a_0 + a_1 X_1 + a_2 X_2$$

and the multiplicative model

$$(2) \quad Y = b_0 X_1^{b_1} X_2^{b_2} .$$

In both models the theoretically correct sign for the X_1 and X_2 parameters is positive. The parameters in both models can be estimated simply by least squares regression if logarithms are taken in (2). Before doing this transformation, however, since inherent risk can be negative it must be transformed. Therefore, the dependent variable in (2) was inherent risk +9, yielding a minimum possible value of +1 (inherent risk can be -8) and hence a minimum logarithm of zero. There can be problems with such a transformation. Results do not stay invariant over such transformations, as pointed out by Goldberg (1971). Hence, the results of this model must be interpreted with caution.

The models were run for all nine product classes and for the data pooled over product classes using dummy variables for the product classes. The results of these runs (coefficients for the dummy variables in the pooled analysis are not shown) are given in Table 3.

Note that the linear models seem to fit slightly better than the multiplicative models. This may be because we have removed the mean and are dealing with only relative risk measures. The parameters have the right sign in all cases where they are significantly different from zero. Finally, note that danger seems to be the more important component in these analyses.

Two additional comparisons with Cunningham's findings can be made. Cunningham used the product classes headache remedies, fabric softener, and dry spaghetti in his study. He found 23.7%, 6.1%, and 1.2% high risk perceivers (by his definition of high risk) for his sample. By using Cunningham's definitions and his ten point certainty and danger measures from this study, 29.9%, 18.6%, and 12.4% of the Ss were classified as high risk perceivers. Thus, Ss in the current study seem to perceive a good deal more risk, by Cunningham's definitions. Finally, Cunningham claims that the certainty and danger components seem relatively independent (a maximum correlation of .22 for his study). However, the correlations between uncertainty and danger for this study range from a low of .294 to a high of .698 across product types, with a correlation of .519 for the data collapsed across product types. Although the correlations for this study will be higher because ten scale points were used as opposed to only four for Cunningham's study, we may conclude that uncertainty and danger are definitely not independent for this study. It is not clear how these differences in findings should be reconciled.

Conclusions

The above findings support to some degree the usefulness of the proposed measures. There are problems, to be sure. For product types with widely dis-

Table 3
Regression Coefficients for Inherent Risk Models

Product Type	Constant	Uncertainty	Danger	R ²
Linear Models				
Paper Towels	-2.816	-.005 (-.008) ^a	.260 (.351)**	.120**
Dry Spaghetti	-.568	.381 (.409)**	.003 (.003)	.169**
Furniture Polish	.740	.358 (.329)**	.017 (.016)	.115**
Toothpaste	.490	.197 (.178)	.322 (.342)**	.208**
Beer	-.006	.153 (.134)	.378 (.342)**	.162**
Instant Coffee	-.640	.036 (.035)	.461 (.450)**	.215**
Aspirin	.349	.323 (.274)*	.403 (.338)**	.319**
Margarine	1.068	.289 (.285)**	.364 (.318)**	.258**
Fabric Softener	-1.991	-.032 (-.035)	.250 (.270)*	.067*
Total ^b	-2.079	.167 (.155)**	.303 (.277)**	.312**
Multiplicative Models				
Paper Towels	1.701	.051 (.117)	.115 (.297)**	.141**
Dry Spaghetti	1.762	.194 (.466)**	-.035 (-.082)	.169**
Furniture Polish	1.787	.164 (.297)*	.009 (.019)	.095**
Toothpaste	1.952	.096 (.169)	.121 (.329)**	.178**
Beer	1.957	.053 (.095)	.133 (.255)*	.087*
Instant Coffee	2.068	-.008 (-.017)	.154 (.367)**	.131**
Aspirin	1.709	.149 (.270)*	.179 (.320)*	.300**
Margarine	1.880	.153 (.342)**	.115 (.243)*	.234**
Fabric Softener	1.959	-.047 (-.104)	.119 (.269)*	.066*
Total ^b	1.656	.082 (.156)**	.116 (.239)**	.285**

a Standardized coefficients (Beta weights) in parentheses. N=97 for product types.

b Analysis for all product classes pooled using dummy variables. The dummy variable coefficients are not given here. N=873 for pooled analysis.

* P < .05

** P < .01

parate risk values (e.g., automobiles and grocery products), it is not clear how to compare the results of different analyses, since we have only relative risk scores. However, the true test of the measure must be based on its usefulness in research or decision making. Studies of the components of risk are in progress using the measures of this study as dependent variables. Other research is needed. In particular, as Spence, Engel, and Blackwell (1970) point out, much more work is needed on the problems of developing criterion measures for risk.

References

- Bauer, R. Consumer Behavior as Risk Taking. In Hancock, R. (ed.), Proceedings of the 43rd Conference of the American Marketing Association, 1960, 389-398.
- Bettman, J. Behavioral Simulation Models in Marketing Systems. Unpublished doctoral dissertation, Yale University, 1969.
- Cox, D. & Rich, S. Perceived Risk and Consumer Decision Making - The Case of Telephone Shopping. Journal of Marketing Research. November 1964, 32-39.
- Cunningham, S. The Major Dimensions of Perceived Risk. In Cox, D. (ed.), Risk Taking and Information Handling in Consumer Behavior. Boston: Harvard Business School, 1967, 82-108.
- Goldberg, L. Five Models of Clinical Judgment: An Empirical Comparison Between Linear and Nonlinear Representations of the Human Inference Process. Organizational Behavior and Human Performance, 1971, 458-79.
- McRae, D. MIKCA: A FORTRAN IV Iterative K-Means Cluster Analysis Program. Mimeo. September 1970.
- Perry, M. & Hamm, B. Canonical Analysis of Relations between Socioeconomic Risk and Personal Influence in Purchase Decisions, Journal of Marketing Research. August 1969, 351-354.
- Sheth, J. & Venkatesan, M. Risk-Reduction Processes in Repetitive Consumer Behavior, Journal of Marketing Research. August 1968, 307-310.
- Spence, H., Engel, J., & Blackwell, R. Perceived Risk in Mail-Order and Retail Store Buying. Journal of Marketing Research. August 1970, 364-369.

RISK ENHANCEMENT AND RISK REDUCTION AS
STRATEGIES FOR HANDLING PERCEIVED RISK

Barbara J. Deering and Jacob Jacoby
Purdue University

Consumers must frequently make purchase decisions on the basis of inadequate information. Accordingly, important decisions often involve consequences which cannot be predicted with confidence. The possible occurrence of unpleasant post-purchase consequences and the pre-purchase uncertainty about those outcomes are incorporated in the concept of perceived risk (Bauer, 1960; Cox, 1967a). According to the risk theoreticians, perceived risk is generally handled through risk-reducing strategies such as brand loyal buying and reliance on certain sources of information (cf. Roselius, 1971). However, risk-enhancement can also constitute a viable risk-handling strategy. Under some circumstances (e.g., boredom), consumers may find the amount of perceived risk unsatisfactorily low and attempt to increase risk through their purchase decisions. A theoretical foundation for understanding risk enhancement and relevant data is now presented.

Perceived risk in a purchase situation depends upon two determinants: the clarity with which the consumer can define decision alternatives, and the consumer's involvement in the decision outcomes (Cox, 1967b). Uncertainty about decision alternatives comes from two sources. First, the consumer's knowledge of her own needs and purchase goals is frequently inadequate. Second, she must usually anticipate purchase results on the basis of those product qualities which she can assess beforehand. Such qualities may or may not have any predictive validity.

Personal involvement refers to the possible adverse consequences of poor purchase decisions. The needs and desires which the purchase is intended to satisfy may remain unaltered. Additionally, the loss of time, money, and energy depletes resources available for future purchases. Severity of perceived consequences thus depends upon the perceived importance of the purchase and the consumer's effort to carry it out.

When the buying situation is defined in terms of perceived risk, it appears to be closely related to Berlyne's (1957, 1960, 1965) concept of conflicting response tendencies. The amount of cognitive conflict evoked varies with three elements: the degree of subjective uncertainty about choosing a response, the absolute strength of the response tendencies, and the degree of incompatibility between responses. The latter two elements correspond closely to constituents of the consequences component in perceived risk. As goals represented by purchase decisions become more important, the consumer would be expected to more strongly desire and attempt to carry out the purchases. The loss of time and money invested are relevant to perceived consequences largely because they are incompatible with the initiation of subsequent responses directed toward the same ends. In most cases, a consumer cannot immediately turn to another purchase alternative if an initial choice has been costly.

Subjective uncertainty is also defined similarly by Berlyne and the risk theorists. According to Berlyne (1965), subjective uncertainty varies with discrepancies in information which confront a person. These discrepancies are named "collative variables", referring to the characteristic comparison,

or collation, of items of information within the present and past situations. Berlyne, like Cox, stresses the subjective evaluation of probabilities (Abelson, Aronson, McGuire, Newcomb, Rosenberg, & Tannenbaum, 1968, p. 45) as the basis for uncertainty.

Also like the risk theorists, Berlyne believes that subjective uncertainty predominates in determining the probability and direction of behavior. However, he asserts that attempts to increase response conflict are as important as attempts to reduce conflict. The function of "diversive" exploration is to introduce variation, new information, "amusement," "diversion," and "aesthetic experience" (Berlyne, 1965, p. 244). Diverive exploration is most likely to occur in monotonous environments. In such environments, a high degree of response conflict will be attractive. Diverive exploration, which does not occur in all circumstances, nonetheless contrasts sharply with the concept of pervasive risk-reducing behavior.

Drawing on Berlyne's formulation, Howard and Sheth (1969) posited two primary purposes for consumers' information search: such search helps fulfill specific purchase goals, either immediately or in the future, and also provides stimulation when the unpredictability of the buying situation is unpleasantly low. This latter, diverive exploration tends to alternate with the former, problem-specific, search as the consumer's attitude toward the product shifts between curiosity and boredom.

Certain evidence suggests that consumers do not act exclusively to minimize the uncertainty and importance of purchases. Buying a brand repeatedly rather than switching involves relatively less uncertainty about product performance. However, repeat buying accounts for only a small proportion of all buying decisions (Frank, 1967). Using available information sources, particularly when the product is new, reduces perceived risk. But consumers confronted with a new brand frequently try it without consulting anyone beforehand (Arndt, 1967). Though mail-order buying generally appears riskier than in-store or salesman-mediated buying, the consumers who shop by mail do not perceive less risk in such buying than do consumers who do not buy by mail (Spence, Engel, & Blackwell, 1970).

Studies dealing specifically with perceived risk offer few data relevant to risk-enhancement for several reasons. First, risk enhancement is more likely under conditions of low perceived risk. Products representing extremely low perceived risk are infrequently used in such studies. They may be eliminated for their apparently trivial and non-involving characteristics--just those characteristics which would be expected to elicit diverive exploration. For instance, after a lengthy series of trials consumers may react to boredom with risk-enhancing behavior. Trial sequences considerably longer than those customarily used (e.g. Sheth & Venkatesan, 1968; Swan, 1969) may be of interest. Second, if product preference and search behavior are expected to both increase and decrease as a function of risk, at least three risk levels are necessary for interpreting risk effects. Use of two risk levels (Cox & Rich, 1964; Spence et al, 1970) or emphasis on a single component of risk (Perry & Hamm, 1969) is relatively uninformative. Third, risk enhancement involves selecting products and items of information which increase the complexity of the consumer's decision. Such diverive stimuli must be available, perhaps in the form of product alternatives, but have not yet been made available in previous studies.

Hypotheses

Given purchase alternatives which encompass a wide range of risk, maximal preference should be manifested for alternatives which are neither extremely high nor low in perceived risk. The most acceptable alternative should represent a compromise between the risk-enhancing and risk-reducing tendencies.

The acceptable range of perceived risk may vary among individuals and groups of individuals, dependent on their previous purchase experience and involvement in unpredictable non-consumer contexts (Howard & Sheth, 1969). Individual differences in risk-taking propensity have been related to willingness to try new products (Popielarz, 1967) and susceptibility to advertising messages (Barach, 1969). Copley and Callom (1971) grouped industrial buyers according to risk perceived across 12 buying situations and found that the relationship of search behavior to perceived risk varied across groups.

Despite individual variability, however, certain decisions probably represent acceptable risk for most consumers. Low cost, easy availability, repetitive use, and low social visibility all decrease the importance and unpredictability of a purchase. Optimal perceived risk would be unlikely to be associated with products extremely low or high in all these attributes. That is, most consumers will prefer purchase alternatives at neither extreme of the perceived risk continuum.

This should be true whether risk is designated in a situation- or object-oriented manner. In the object-oriented approach (Sandell, 1968), individuals evaluate particular products in terms of the associated unpredictability and importance. The situation-oriented approach recognizes that certain situations can be expected to alter uncertainty and/or importance across products. For instance, introduction of a new brand represents risk increase because the consumer's knowledge becomes incomplete to some degree. The addition of a guarantee usually increases the consumer's information regarding the product, thus representing a form of risk reduction.

If, as the risk-optimization process suggests, perceived risk and preference are positively related in purchase decisions involving extremely low risk and negatively related in decisions involving extremely high risk, then interaction between product and situation risk is implicit. Either the product or situation can be conceptualized as the determinant of the basic level of risk involved in a decision, with the other risk element providing increases and decreases relative to that level. For instance, if product risk is considered basic, then a positive relationship between situation risk and preference would be expected at low levels of product risk and a negative relationship at higher levels. Conversely, negative and positive relationships between preference and product risk would be expected in situations high and low in risk, respectively.

The following study examined preference as a function of perceived risk associated with both products and situations. The specific hypotheses entertained were: (1) purchase preference is curvilinearly related (inverted-U) to the amount of risk associated with a set of ordered product alternatives; (2) this relationship occurs whether risk is determined by the products or the information provided in the purchase situation; (3) product and situational determinants interact to affect preference.

Method

Overview

A written questionnaire which described three shopping situations was administered to two independent samples of female homemakers. Twenty products remained constant across these three situations, and ten questions measured the theoretical components of perceived risk associated with each of these products. Respondents role-played to indicate how they would react in each situation. The 20 products were selected from 56 pre-tested products and represented a wide total range of perceived risk with relatively small dispersion among judgments of risk associated with any single product.

Subjects

Each sample was randomly selected from a list of residences in each of two cities. Sample 1 was drawn from Richardson, Texas, a suburb of Dallas with a population of approximately 50,000. Sample 2 was drawn from Lafayette and West Lafayette, Indiana, two non-suburban cities with a combined population of approximately 100,000. In Sample 1, 118 respondents, or 88% of those contacted, returned questionnaires. In Sample 2, 111 returns represented 96% of the number distributed. Respondents' failure to follow instructions necessitated eliminating two questionnaires in each group.

Administration

Questionnaires were personally delivered to respondents' homes and retrieved three days later. The deliverer described the study as an academically-sponsored project, involving objective questions about consumer behavior. Each person who completed a questionnaire later received \$2.00 in payment and a letter explaining the purpose of the study.

Measures of Product Risk

Ten questions (see Table 1) were constructed to assess components of perceived risk for 20 products. These items were responded to on a 9-point scale in which high values on each scale indicated a high degree of danger or uncertainty. The order of products was randomized across questions; question order was randomized across subjects. Responses to selected questions were combined to form three composite measures. As defined by Cox (1967b, 1967c), perceived risk requires the combination of two general risk components; purchase-relevant uncertainty and consequences. Each risk component also has several dimensions. For comparison, each of the three composite measures emphasized somewhat different dimensions of the uncertainty and consequences components.

The first composite measure (CM-1) combined responses to two questions used in previous studies (Cunningham, 1967a, 1967b) and labeled A and B in Table 1. CM-1 was obtained by multiplying the individual's scores on these items, yielding a score range from 1-81.

In the second composite measure (CM-2), the uncertainty components emphasized individual-specific differences in ability to predict product attributes. Questions C, D, and E were combined as follows to give a measure with a range of 1-81: $CM-2 = (C) \cdot (D + D)/2$.

In the third composite measure (CM-3) of perceived risk, consequences were again represented by ratings of importance (D) and investment (B), as in CM-2.

The unpredictability component included the perceived unpredictability of product dependability with repeated use (F), product construction and materials (G), results of product failure (H), and the degree (I) and kind (J) of goal fulfillment involved. The formula for obtaining CM-3 was: $CM-3 = (D + E)/2$. $(F + G + H + I + J)/5$.

Table 1

Questions Measuring Perceived Risk

-
-
- A. How certain are you that a brand name of this product you haven't tried will work as well as your present brand?
- B. We all know that not all products work as well as others; compared to other products, how much danger would you say there is in trying a brand of this product that you have never used before?
- C. How confident would you say you are about judging the quality of the product?
- D. Buying a product that gives you good results may be more important for some products listed than for others. How important would you say it is for this product to satisfy you?
- E. The investment you make when you buy a product includes your time and energy as well as money. In terms of the time, money, and overall effort required to buy this product, how much would you say you invest?
- F. Can most shoppers guess ahead of time how dependable this product will be if it is used over and over again?
- G. Before buying this product, can almost anyone tell how good its materials are and how well it's put together?
- H. Can almost any shopper predict what the bad results will be if this product fails?
- I. In general, does this product tend to fulfill your expectations?
- J. Is it obvious why someone like yourself would want this product?
-

Measures of Situational Risk

Each respondent was given the opportunity to win one of the 20 products, except the four most expensive, as a prize in a drawing to be made from every 20 questionnaires. Respondents indicated on a 9-point scale how much they would like to receive each gift as a prize. Subsequent questions, ordered alternately across subjects, assessed product preferences in two situations: (1) products were offered with a full money-back guarantee against "any failure to perform satisfactorily during normal use"; (2) products were described as new brands, soon to be marketed for the first time. The guaranteed and new products represented, respectively, risk-reducing and risk-enhancing circumstances.

Results

Composite Risk Measures

Products were ranked according to each of the composite measures of risk.

Rankings of the two samples were closely related: the Spearman rank-order correlation between the two samples of CM-1 equalled .9323, and was high also for CM-2 ($r_s = .9097$) and CM-3 ($r_s = .9301$).

The mean of the two sample ranks for each product was calculated for each composite risk measure. The mean of the two sample ranks for each product was similar across the three risk measures: r_s for CM-1 and CM-2 was .8925; for CM-1 and CM-3, $r_s = .8684$; for CM-2 and CM-3, which used the same measure for the consequences component, $r_s = .8820$. The mean of the three risk measures was then computed for each product and used to classify products into one of five perceived risk levels. Each level was relatively homogeneous in terms of numerical indices of perceived risk for the products included in that level. The final mean product ranks (the cross-measure mean of cross-group ranks) and corresponding risk measures appear in Table 2.

Table 2

Rank, Mean Perceived Risk, and Final Risk Level of 20 Products in Combined Samples

<u>Rank</u>	<u>Product</u>	<u>Mean Risk</u>	<u>Risk Level</u>	<u>Mean of Level</u>
1	bath towels	10.0868	1	11.6640
2	dry spaghetti	11.1137	1	
3	salt	11.7288	1	
4	raincoat	13.7265	1	
5	ice cream	14.2125	2	14.9530
6	saucepan	14.7275	2	
7	table lamp	14.6733	2	
8	lingerie	16.1988	2	
9	fabric softener	18.5293	3	20.3646
10	dress shoes	21.0791	3	
11	beer	19.8799	3	
12	record player	21.9702	3	
13	gasoline	22.5577	4	23.9498
14	detergent	23.0688	4	
15	floor wax	24.2898	4	
16	refrigerator	25.8827	4	
17	headache remedy	29.2855	5	31.3874
18	hair spray	26.6924	5	
19	automobile	32.8151	5	
20	car tires	36.7564	5	

1 = low risk; 5 = high risk

Analysis of Product Preferences

Product preferences provided cell entry data for a 2 (Samples) X 3 (Situation Risk) X 5 (Risk Level) analysis of variance, with repeated measures on the second and third factors. The two samples were treated as the first, between-subject factor. An unweighted means solution was applied for unequal sample sizes (Winer, 1962, pp. 374-378).

The model for a repeated-measures analysis is appropriate if the variance-covariance matrices are equal and the pooled variance-covariance matrix is

symmetrical (Winer, 1962, pp. 369-374). Each portion of the within-cell variation was tested for homogeneity of variance, using Hartley's F_{\max} test. In-

significant results were obtained for variation due to interaction between subjects within groups and situation risk ($F_{\max} .95 [2,223]$), product risk

($F_{\max} .95[2,546]$), or both product and situation risk ($F_{\max} .95[2,892]$). Complete assumptions about the variance-covariance matrices were not tested.

Consequently, the analyses used conservative degrees of freedom (Greenhouse & Geisser, 1959). The conservative procedure is negatively biased, i.e., the critical value is over-large, leading to errors in the direction of not rejecting false null hypotheses.

Results of the 2 X 3 X 5 analysis of variance are summarized in Table 3.

Table 3

Analysis of Variance of Product Preference for Products, Situations, and Samples

Source	df	MS	F
Between subjects:			
A (Sample)	1	11.831	.124
A X Subjects within groups	223	95.590	
Within subjects:			
B (Product)	4	2173.561	184.655*
A X B	4	143.733	12.211*
B X Subjects within groups	892	11.771	
C (Situation)	2	2281.368	47.257*
A X C	2	156.874	3.250
C X Subjects within groups	446	48.276	
B X C	8	213.585	23.186*
A X B X C	8	218.798	23.752*
B X C X Subjects within groups	1784	9.212	

* $p < .0005$, $df = 1,225$

Sample main₂ effect was insignificant; variation within risk treatments accounted for most ($\omega^2 = .78$) of the total response variation. Situation and product risk factors each produced significant main effects (respectively, $F = 184.655$, $df = 1,225$, $p < .0005$; $F = 47.257$, $df = 1,225$, $p < .0005$). According to Newman-Keuls analysis of means representing main effects, preference decreased significantly ($p < .01$) between each of the first three product levels and changed insignificantly at higher levels. Preference in the guarantee situation was

significantly ($p < .01$) greater than in the other two situations, which did not differ significantly.

As expected, both aspects of risk contributed to product preference. Interaction between product and situation risk factors was significant ($F = 23.186$, $df = 1,225$, $p < .0005$). The interaction (see Figure 1) was evaluated in terms of the effect of one risk factor at each single level of the other risk factor. Analysis of variance of simple product risk effects yielded significant results for the guarantee ($F = 104.145$, $df = 1,225$, $p < .0005$), free gift ($F = 100.497$, $df = 1,225$, $p < .0005$), and new brand situations ($F = 268.947$, $df = 1,225$, $p < .0005$). Within each of the latter two situations, mean preference decreased significantly ($p < .01$) between product risk levels 1 and 2, and levels 2 and 3, according to Newman-Keuls tests. At levels 3 and 4, means of preference for a free gift did not differ, and preference increased ($p < .01$) at level 5. Preference for a new brand increased ($p < .01$) between levels 3 and 4 and decreased ($p < .01$) at level 5. Within the guarantee situation, mean preference was highest for products at the lowest product risk level. Preference decreased at product level 2 and remained stable across levels 3, 4, and 5.

Analysis of simple effects of situation risk indicated significant ($df = 1,225$, $p < .0005$) effects at each level of product risk. All inter-mean comparisons of preference across situations at each level of product risk were significant ($p < .05$) in Newman-Keuls tests, with one exception. Preference at product risk level 2 differed insignificantly for the guaranteed and new products.

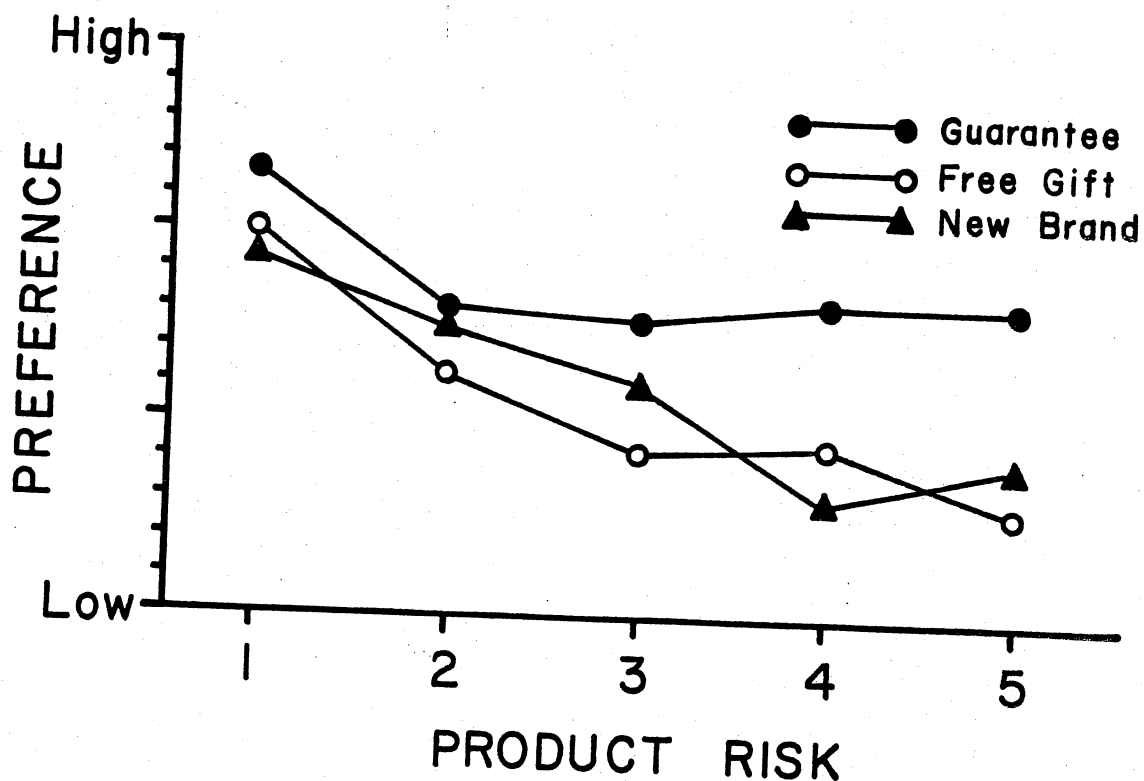


Fig. 1. Interaction between product and situational risk.

The hypothesized attractiveness of an optimal amount of risk implies higher preference at one risk level and lower preference at adjacent levels. Given a comprehensive range of risk levels, curvilinear relationship of preference to risk levels would be expected. Accordingly, trend analyses were made for the two risk main effects. Orthogonal coefficients were derived for unequal product risk intervals and the unequal sample sizes (Kirk, 1968, pp. 513-517). Situational risk intervals were assumed to be equal. For product risk, the linear, quadratic, and cubic trend components were significant ($F = 480.6788$, $df = 1,56$, $p < .0005$; $F = 142.5231$, $df = 1,56$, $p < .0005$; $F = 11.4949$, $df = 1,56$, $p < .0005$). Contrary to expectation, the linear trend clearly dominated, contributing 85% of the variance. The trend of situational main effects was predominantly quadratic ($\omega^2 = .87$), although both linear and quadratic trend components were significant ($F = 822.9858$, $df = 1,112$, $p < .005$; $F = 136.5222$, $df = 1,112$, $p < .0005$).

Trend analyses were conducted to compare the pattern of effects of one risk factor at each level of the other factor. Linear, quadratic, cubic, and quartic components of simple product risk effects were significant ($df = 1,56$, $p < .005$) at each level of situation risk. The quadratic component of product risk accounted for more of the treatment variance ($\omega^2 = .55$) in the guarantee situation. In the gift and new brand conditions, the linear trend component dominated ($\omega^2 = .66$; $\omega^2 = .89$, respectively). The relative contribution of the quadratic component decreased markedly between the gift ($\omega^2 = .28$) and the new brand conditions ($\omega^2 = .01$). The pattern of product risk effects thus tended to differ at different levels of situational risk.

Trend analysis of simple situation risk effects produced significant ($df = 1,112$, $p < .005$) linear and quadratic components at all levels of product risk. The linear trend component dominated at the lowest product risk level ($\omega^2 = .88$). Quadratic trend accounted for more treatment variance at levels 2 and 3 ($\omega^2 = .76$; $\omega^2 = .85$, respectively). At product risk levels 4 and 5, linear and quadratic components contributed equally ($\omega^2 = .49$; $\omega^2 = .51$, at level 4; $\omega^2 = .55$ and $\omega^2 = .45$ at level 5).

In the 2 X 3 X 5 analysis, all interactions involving the samples were significant (see Table 2). To clarify the effect of risk factors within each sample, simple main effects for each risk factor, and simple interactive effects for the product-situation combination were calculated. Simple product risk effects were significant for both Sample 1 ($F = 106.044$, $df = 1,116$, $p < .0005$) and Sample 2 ($F = 158.561$, $df = 1,109$, $p < .0005$), as were simple situation risk effects (for Sample 1, $F = 63.334$, $df = 1,116$, $p < .0005$; for Sample 2, $F = 60.409$, $df = 1,109$, $p < .0005$). Simple product-situation interaction occurred for Samples 1 and 2 (respectively, $F = 10.065$, $df = 1,116$, $p < .001$; $F = 41.480$, $df = 1,109$, $p < .0005$). Situation and product risk and their interaction influenced product preference within as well as across the two samples. Additionally, trend analysis of the simple main risk effects yielded no notable deviations from the trends reported for the combined samples. The trends for product and situational risk effects were attributable primarily to linear and quadratic components, respectively, within each sample.

Although the effect of each risk factor was similar across samples, interaction between the risk factors varied somewhat with samples (Figure 2).

Analysis of variance of simple interaction between samples and product risk indicated significant interaction in the gift ($F = 12.460$, $df = 1,225$, $p < .001$), and new brand situations ($F = 12.460$, $df = 1,225$, $p < .001$). For the latter situation, Newman-Keuls analysis indicated significantly ($p < .05$) greater preference for Sample 1 for product risk levels 2 and 3, and significantly ($p < .001$) less preference at product risk level 5. A single significant inter-mean difference occurred in the free gift situation. Preference for Sample 2 was significantly ($p < .05$) greater at product risk level 3. The preceding significant inter-mean differences were sufficient to give significant simple interaction between samples and situational risk at product risk levels 2, 3, 4, and 5.

Discussion

Consumers' product selections provided strong support for the hypothesized interaction between situational and product determinants (Hypothesis 3). Additionally, each determinant alone also affected product preference, as posited by Hypothesis 2. The latter hypothesis also stipulated a curvilinear relationship between preference and the continuum of risk associated with a set of ordered alternatives. However, the data indicated that preference tended to decrease with increasing product risk and then to stabilize at the higher risk levels. Similarly, the guarantee situation elicited the greatest preference, but preference did not differ between the well-known and new brand situations.

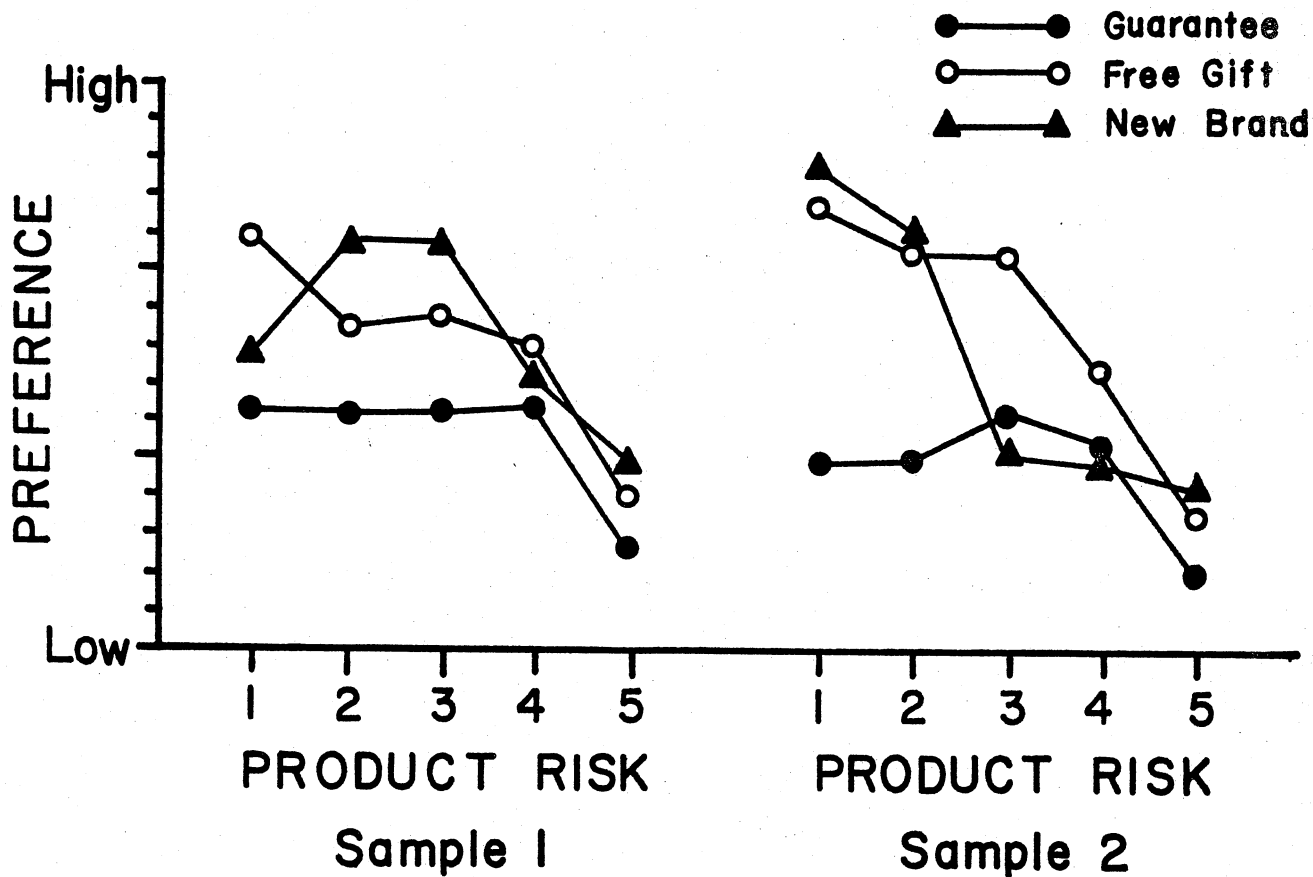


Fig. 2. Interaction of situation risk, product risk, and sample.

Although the predicted curvilinear relationship did not occur for situational and product determinants separately, this does not preclude the possibility of the predicted relationship occurring for combined situational and product risk. However, as joint determinants, product and situation risk interacted to elicit greatest preference when both risk components were small. Preference decreased slightly as both risk components increased, with the rate of decrease being greater for the higher risk situations. In contrast, a risk-optimization process should produce increases and subsequent decreases in preference (i.e., an inverted-U function) with increases in perceived risk.

Despite the apparent simplicity of the overall interaction between risk components, the underlying relationships were complex. Within the "safest" (i.e., guaranteed products) situation, preference stabilized at a low level of product risk. In the intermediate risk (i.e., well-known brand) situation, product risk and preference showed a consistent negative relationship. And within the riskiest (i.e., new brand) situation, important sample differences emerged. One sample behaved much as it had within intermediate risk. In the other sample, the relationship between preference and perceived risk tended toward a U-shaped pattern.

Sample differences were one of the more interesting findings of the study. One sample of consumers responded to greater situational risk with decreases in preference, except when the products involved represented very high risk. With these products, preference was lowest in the intermediate risk situation and high when either a guaranteed or new brand of product was offered. The second sample provided this same U-shaped function for situation risk when products were intermediate (levels 2 and 3) in risk and a negative risk-preference relationship for other products. It is notable that risk enhancement in each case involved selection of a new product brand rather than particular products representing higher risk.

To clarify sample differences, several measures of search for information and of emotional response to purchase situations were examined for each sample. The sample with risk-enhancing preferences (e.g., selection of new products) at the highest product level experienced significantly less worry, relative to the other sample, when shopping for products at that level. In contrast, risk-enhancing preferences at product level 3 were associated with less desire for increased product variety at that level. It appears that those who enhanced risk at a particular level were relatively more satisfied, in terms of the variety offered or associated worry, with risk at that level. Additionally, those who enhanced risk at a lower product level seemed to attribute greater risk to the risk-enhancing situation. They were more likely to read and deliberate about purchasing a new brand. Despite the specification of product and situational risk factors, individual and group differences in interpreting product characteristics persisted.

Identifying consumers who would enhance risk at a relatively low or high risk level would be useful in anticipating market response to product changes. In the present study, age information was available for homemakers in the samples. Those who enhanced risk at a relatively low level were equally represented across all intervals from 20 to 80 years. Risk enhancement at higher levels was more likely when the homemakers were between 20 and 40 years of age. Since these years tend to require the greatest homemaking activity, this may accustom the consumer to a higher level of risk. Attempts to enhance risk may center about that higher level.

That risk enhancement occurred at certain risk levels and for certain groups suggests two areas for future research. First, it is possible that risk enhancement centers on a range of perceived risk, with risk reduction behavior occurring both below and above that range. This implies that not all products contribute equally to the risk-optimization process. With products representing very low risk, the effort involved in selecting new products may not be justified by the transient risk enhancement.

The second area concerns identification of groups who respond similarly to perceived risk. Self-confidence, both product-specific and general, is one psychological variable that has been used with some success (Barach, 1969). Although other variables such as cognitive complexity, category width, and tolerance for ambiguity should be examined, certain experiential measures may be equally helpful. Making many purchases perceived as high in risk, or a relatively high level of shopping and other homemaking activities, should, over time, accustom a consumer to a relatively high level of perceived risk. The amount of risk perceived in a purchase situation would be established relative to this higher norm.

For individuals in both samples, the norm of risk experienced in actual shopping activities may have obscured the relationship between perceived risk and product preference. The highest obtained average value for perceived risk ($\bar{X} = 36.76$) was less than half the range on the scales provided. Purchases incorporating greater perceived risk may occur. For example, the monetary commitment required for home furnishings or insurance probably exceeds that of any of the products studied, and social implications may be equally as great. If such purchases were included for consideration, it is possible that they would elicit the decrease in preference predicted for extremely high risk. The generally negative relationship found between risk and preference may thus reflect an insufficient range of perceived risk.

An accurate a priori estimate of the range of risk associated with products may be difficult. Highest perceived risk was associated with purchasing a car, car tires, hair spray, and a headache remedy. Risk apparently derives from diverse consequences of product use. Jacoby and Kaplan (1972) detected five varieties of perceived risk: functional, physical, social, financial, and psychological. Relative contribution of each variety to overall perceived risk varied across products. Such diversity in components of risk may affect the relationship between risk and product preference or search.

Product preference is one of many ways of handling perceived risk (Roselius, 1971). Both the relevance of information to particular product characteristics and the sources of information may vary with perceived risk. Activities which precede and follow product choice must also be examined to evaluate the possibility of risk optimization.

References

- Abelson, R. P., Aronson, E., McGuire, W. M., Newcomb, T. M., Rosenberg, M. J., & Tannenbaum, P. H. (Eds.), Theories of cognitive consistency: A source book. Chicago: Rand-McNally, 1968.
- Arndt, J. Perceived risk, sociometric integration, and word-of-mouth communication in the adoption of a new food product. In Donald F. Cox (Ed.), Risk-taking and information-handling in consumer behavior. Boston: Harvard University Press, 1967.

- Barach, J. A. Advertising effectiveness and risk in consumer decision process. Journal of Marketing Research, 1969, 6, 314-333.
- Bauer, R. A. Consumer behavior as risk taking. In R. S. Hancock (Ed.), Dynamic marketing for a changing world. Chicago: American Marketing Association, 1960.
- Berlyne, D. E. Uncertainty and conflict: A point of contact between information theory and behavior theory concepts. Psychological Review, 1957, 64, 329-339.
- Berlyne, D. E. Conflict, arousal, and curiosity. New York: McGraw-Hill, 1960.
- Berlyne, D. E. Structure and direction in thinking. New York: Wiley, 1965.
- Copley, T. P., & Callom, F. L. Industrial search behavior and perceived risk. Proceedings, Second Annual Convention of the Association for Consumer Research, 1971, 208-31.
- Cox, D. F. (Ed.) Risk-taking and information-handling in consumer behavior. Boston: Harvard University Press, 1967. (a)
- Cox, D. F. Introduction. In D. F. Cox (Ed.), Risk-taking and information-handling in consumer behavior. Boston: Harvard University Press, 1967. (b)
- Cox, D. F. Risk handling in consumer behavior--Intensive study of two cases. In D. F. Cox (Ed.), Risk-taking and information-handling in consumer behavior. Boston: Harvard University Press, 1967. (c)
- Cox, D. F., & Rich, S. U. Perceived risk and consumer decision making--The case of telephone shopping. Journal of Marketing Research, 1964, 1, 32-39.
- Cunningham, S. M. The major dimensions of perceived risk. In D. F. Cox (Ed.), Risk-taking and information-handling in consumer behavior. Boston: Harvard University Press, 1967. (a)
- Cunningham, S. M. Perceived risk as a factor in informal consumer communication. In D. F. Cox (Ed.), Risk-taking and information-handling in consumer behavior. Boston: Harvard University Press, 1967. (b)
- Frank, R. E. Is brand loyalty a useful basis for market segmentation? Journal of Advertising Research, 1967, 7, 27-33.
- Greenhouse, S. W., & Geisser, S. On methods in the analysis of profile data. Psychometrika, 1959, 24, 95-112.
- Howard, J. A., & Sheth, J. N. The theory of buyer behavior. New York: John Wiley, 1969.
- Jacoby, J., & Kaplan, L. B. The varieties of perceived risk. Purdue University, Purdue Papers in Consumer Psychology. Paper No. 118, 1972.
- Kirk, R. E. Experimental design: Procedures for the behavioral sciences. Belmont, California: Brooks/Cole Publishing, 1968.
- Perry, M., & Hamm, B. C. Canonical analysis of relations between sociometric risk and personal influence in purchase decisions. Journal of Marketing Research, 1969, 6, 351-354.
- Popielarz, D. T. An exploration of perceived risk and willingness to try new products. Journal of Marketing Research, 1967, 4, 368-373.
- Roselius, T. Consumer rankings of risk reduction methods. Journal of Marketing, 1971, 35, 56-61.
- Sandell, R. G. Effects of attitudinal and situational factors on reported choice behavior. Journal of Marketing Research, 1968, 5, 405-408.
- Sheth, J. H., & Venkatesan, M. Risk reduction processes in repetitive consumer behavior. Journal of Marketing Research, 1968, 5, 307-310.
- Spence, H. E., Engel, J. F., & Blackwell, R. D. Perceived risk in mail-order and retail store buying. Journal of Marketing Research, 1970, 7, 364-369.
- Swan, J. E. Experimental analysis of predecision information seeking. Journal of Marketing Research, 1969, 6, 192-197.
- Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

CONSUMER REACTION TO
RESTRICTION OF CHOICE ALTERNATIVES¹

Michael B. Mazis
and
Robert B. Settle²
University of Florida

On January 1, 1972, Dade County Florida began prohibiting the sale, possession, or use of any detergent or cleaning agent containing phosphates. For the next several weeks, Miami housewives who visited their local supermarket to purchase laundry detergent found that their choice alternatives were diminished from a few dozen brands to two or three, and in some cases there was no choice whatsoever. There appeared to be three basic responses which the consumer could make to this diminished choice situation: (1) she might react negatively toward the new, no-phosphate detergent, (2) she might prefer the new detergent to the phosphate detergent, or (3) she might be indifferent to the old and the new, regardless of the diminished choice.

The dramatic changes in the consumer's choice alternatives created by the legal restraints against phosphates appeared to provide an opportunity to test the theory of psychological reactance in the field situation, and in relation to a significant marketing problem. The formulation and investigation of psychological reactance has been associated principally with experimental research and in the laboratory, and this association has necessarily created some questions concerning the external validity and generalizability of the concept. A study of consumer reaction to real choice limitations in the market appeared appropriate.

Theoretical Framework

The theory of psychological reactance, as formulated by Brehm (1966) states that it is very important to the individual to maintain his choice alternatives. If he is able to do so, he will be better able to maximize the rewards he obtains from his behavior. As a consequence of this association between the number of alternatives available and the reward value of action, the individual tends to prefer a wide variety of alternatives, and learns to protect these options even under circumstances where he may not wish to exercise them. Thus, the diminishment of a person's choice alternatives will lead to a motivational state such that the person will try to reinstate the lost alternatives; he will respond in a reactive manner.

Several examples of reactive behavior are easily available to the casual observer, and they make the concept intuitively clear. The small child, told to avoid the oven door because it is hot, will often proceed immediately to put his hand on the stove. The woman who enters an ice cream shop and is told that guava and rhubarb sherbert is not available today might be heard to mutter about today being the day she was going to try it. In the extreme situation, a person who is forced to accept the alternative he would otherwise have selected may try to refuse it and attempt to obtain an alternative he would never have chosen under a free choice situation. All are examples of reactive behavior.

Antecedent and Consequent Conditions of Reactivity

The degree of reactance experienced by an individual is affected by a variety of factors. Principal among them are: (1) the strength of the rewards, mediated by the behavior alternatives, (2) the number of alternatives eliminated and the number remaining, (3) the perceived power and intent of the force which limits the choice alternatives, and (4) the perceived threat to the remaining alternatives and to those in other choice situations. In addition to these factors, the propensity to be reactive might prove to be an individual difference variable operative over a wide variety of situations.

While the factors listed constitute the antecedent conditions of a motivational state termed reactance, there are several behavioral possibilities open to the individual which can be recognized as the consequent conditions. The most common modes of response to reactance are: (1) to attempt to re-establish the alternatives which were removed, (2) to re-evaluated upward the value of forgone alternatives and re-evaluate downward the value of the remaining alternatives, (3) to exercise alternatives in other choice situations parallel to those which were removed, (4) to downgrade the "necessity" or reason for the choice limitation, (5) to manifest increased negativity and opposition to the entity which removed the choice alternatives, and (6) to encourage other people to exercise the denied choice alternatives.

To the extent that the antecedent conditions of reactance are present in the Miami area after the ban on phosphate detergents, the shopper can be expected to react along the lines described as consequent conditions, if reactance theory correctly describes the psychological process associated with a diminishment of choice. Other psychological theory, however, suggests rather different modes of response. The theory of cognitive consistency predicts that the shopper in this situation would tend to accept the new no-phosphate detergent and perhaps prefer it because she has learned of the contribution of phosphates to pollution. The law might be seen as an input of new information, and the negative pronouncements concerning phosphates would lead the consumer to down-grade the old and up-grade the new detergent. Similarly, she might exaggerate the threat created by phosphates and to enhance her positive opinion of the authority which has removed the phosphate detergent from the market.

For a variety of reasons, the consumer may manifest behavior which supports neither psychological theory. For example, she may perceive little difference between phosphate and no-phosphate detergent because she perceives few differences among brands of detergent, because the item is not particularly important or expensive, or because she is not ego-involved in a socially invisible good. Any such reason for detachment from this situation may lead to a failure to discriminate.

The Field Survey

A five part questionnaire was constructed, with the first section designed to obtain the necessary information on the brand and type of detergent being used by the respondent presently and in the past. If the respondent indicated use of a no-phosphate detergent, she was also asked why she had switched, to measure whether or not she perceived herself as

having a choice. This section also obtained the respondent's opinion concerning the amount of no-phosphate detergent and amount of extra ingredients used, compared to that used with the old detergent.

The second section of the questionnaire obtained the respondent's ratings of the phosphate and no-phosphate detergent on seven evaluative criteria, together with an overall rating of each. Ratings were made separately, and no comparison was solicited or suggested. Obviously, only those respondents using no-phosphate detergent were asked to rate this product. Respondents were also asked in this section what products were banned in Dade County, to measure their awareness of the law. They were also asked their feelings concerning the law banning the sale or use of detergents containing phosphates. Lastly, this section of the questionnaire obtained information about stocking up (hoarding) and about the purchases outside of Dade County (smuggling).

Section three of the questionnaire dealt with the respondent's attitudes toward: (1) water pollution, (2) the contribution of phosphates to pollution, (3) the role of the government in protection from pollution, and (4) legal restrictions against the sale of phosphate detergent. Respondents were given a card containing a series of five statements on these topics and asked to indicate if they strongly agreed, agreed, were neutral, disagreed, or strongly disagreed with each. A second card contained four statements concerning attitudes toward washing, and the respondent answered in a similar manner.

The fourth section of the questionnaire solicited information about the respondent's values and personality. First she was given a card containing five social goals or objectives: equality, freedom, family security, a clean environment, and social recognition, adopted from Rokeach's Value Survey (1968). Each was followed by a short explanatory phrase, and the respondent was asked to rank them in their order of importance to her. Second, she was given a card containing five pairs of statements taken from Rotter Internal-External Locus of Control Scale (1966). Each pair contained one statement identifying an internal attribution and one statement of an external attribution. Respondents were asked to indicate which of each pair they believed most strongly.

The last section of the questionnaire was concerned with the respondents judgement of two pieces of terrycloth material. The two were identical, however, one was clearly labeled "phosphate" and the other, "no-phosphate." The respondent was told that both had been washed several times, each in a different kind of detergent. She was asked to judge which was cleaner, which was softer, and which was fresher. Finally, the necessary demographic data was obtained.

Administration

The data were obtained in personal interviews in the subjects' homes. The four female interviewers employed in Miami and the two in Tampa completed 84 and 45 interviews, respectively. The interviewers were instructed in the use of the questionnaire and the nature of the study, but remained ignorant of the specific hypotheses. They were permitted to select a judgement sample within constraints on age, sex, family, geographic area, dwelling area status, and house type. The data were collected within eight weeks after the law became effective.

Results and Conclusions

This report focuses on the patterns of interaction between three main elements: area of residence (Miami versus Tampa), respondent attitudes, and perceptual distortion. In addition to these, differential patterns of response are examined between those in the Miami group who were required to switch brands because their favorite brand was no longer available and those who were able to maintain brand continuity. The relationship between the factors cited above and the two personality variables, value orientation and locus of control, were also investigated; however, these did not prove to be systematically related. Several elements did prove sufficiently sensitive, however, to test the basic propositions of the theory of psychological reactance.

Attitudinal Differences

The theory of psychological reactance suggests that limitations in choice alternatives might lead to a motivational state to re-establish the alternatives. If this is not possible, one would expect the frustration to result in attitude change. Table 1 depicts an index of attitudes toward washing clothes resulting from answers to four questions, for respondents in Miami and those in Tampa. Only a third of Miami residents maintained positive attitudes toward washing, while nearly twice that proportion in Tampa were favorable. The difference in the distribution proved to be significant

Table 1

Index of Attitudes Toward Washing, By Area Group

Area	N	Positive	Negative
Miami	49	31%	69%
Tampa	31	61%	39%

Chi square = 6.11 df = 1 p < .02

The attitude of respondents toward government protection of clean water is displayed in Table 2. While only five percent of Tampa residents were hostile toward the government's role, a quarter of Miami residents so indicated. Reactance theory suggests that negativity toward the source of choice deprivation is a consequent condition of the motivational state, thus these results tend to support a reactance interpretation.

Table 2
Attitude Toward Government Role,^a
by Area Group

Area	N	Agree	Disagree
Miami	71	75%	25%
Tampa	44	95%	5%

Chi square = 6.80 df = 1 p < .01

^aThe government should play an important role in protecting our water from pollution.

The attitude of respondents toward legal restrictions against phosphate detergents are depicted in Table 3. Again, those in the choice deprivation condition manifested significantly more negative attitudes, supporting the hypothesis that they have experienced reactance.

Table 3
Attitude Toward Legal Restrictions,^a
by Area Group

Area	N	Agree	Disagree
Miami	71	63%	37%
Tampa	35	89%	11%

Chi square = 6.14 df = 1 p < .02

^aLegal restrictions should be imposed against the sale of detergents containing phosphates.

Perceptual Distortion

Tables 4, 5, and 6 contain the results of the questions concerning which piece of material was cleaner, softer, and fresher. One was labeled as having been laundered in phosphate detergent and the other in non-phosphate detergent.

Table 4 indicates the distributions of response are significantly different, and the Miami group shows a higher proportion selecting the no-phosphate material as cleaner. Since this result, if different from chance, would be antithetical to a reactance explanation, it was examined further. The chi square value for the distribution of response without the "same" category did not prove to be significant, and consequently it can be concluded that the differences in the distributions result from the fact that more than twice as many of the Tampa respondents felt both pieces of material were equally clean.

Table 4
Material Judged Cleaner, By Area Group

Area	N	Phosphate	No Phosphate	Same
Miami	84	33%	53%	14%
Tampa	45	27%	38%	35%

Chi square = 7.85 df = 2 p < .05

Table 5 also indicates significantly different distributions of response between Miami and Tampa, however, in this case, where respondents judged the softness of the two pieces of identical material, Miami residents chose the phosphate detergent product in a markedly higher proportion. When the "same" category was eliminated, and a two-by-two chi square was obtained, it proved to be significantly different beyond the .01 level. Thus, a significantly larger proportion of Miami residents selected the phosphate detergent product as superior, and this constitutes support for the reactance theory explanation.

Table 5
Material Judged Softer, By Area Group

Area	N	Phosphate	No Phosphate	Same
Miami	84	57%	31%	12%
Tampa	45	20%	44%	36%

Chi square = 18.7 df = 2 p < .01

The freshness of the material was judged in the third question of this section, and the results are shown in Table 6. The distributions are significantly different with only marginal certainty, however, the table indicates that again a larger proportion of Miami respondents perceived a difference in the identical pieces of material. It appears that the Miami group, having been exposed to information concerning the relative effectiveness of the two types of detergent, were more likely to discriminate between detergents. In one case, for the judgement about softness, respondents were significantly more favorable toward the phosphate detergent, as reactance theory would predict.

Table 6
Material Judged Fresher, By Area Group

Area	N	Phosphate	No Phosphate	Same
Miami	84	27%	39%	23%
Tampa	45	24%	33%	42%

Chi square = 5.74 df = 2 p < .10

Tables 7 and 8 contain the data on perceptual distortion cross-tabulated with the respondents' attitudes on the danger of phosphate pollution. Table 7 contains the distributions of response on which material was judged to be cleaner. A higher percentage of those who felt that phosphates did not constitute an important threat also perceived the phosphate detergent-laundered material as being cleaner. Table 8 depicts the distributions on the material judged softer, according to the same attitude. Nearly all those who agreed phosphates were not an important threat also perceived the phosphate-laundered material as softer. Neither result constitutes a test of reactance, however, both demonstrate the consistency of attitudes and perceptual distortion.

Table 7
Material Judged Cleaner, by Attitude Toward Phosphate^a

Attitude	N	Phosphate	No Phosphate
Agree	19	53%	47%
Disagree	33	24%	76%

Chi square = 3.13 df = 1 p < .10

^aPhosphates from detergents do not constitute an important threat compared to other forms of water pollution.

Table 8

Material Judged Softer, by Attitude Toward Phosphate^a

Attitude	N	Phosphate	No Phosphate
Agree	18	94%	6%
Disagree	32	56%	44%

Chi square = 6.29 df = 1 p < .02

^aPhosphates from detergent do not constitute an important threat compared to other forms of water pollution.

Brand Continuity

Within the Miami experimental group, nine of the 86 respondents continued to use phosphate detergents when the survey was taken. These individuals either hoarded the product or smuggled it in from a nearby county. Thus, over ten percent of those interviewed can be regarded as highly reactive, since they risked possible prosecution to maintain their choice alternatives.

Of those respondents from Miami using no-phosphate detergent, some found it necessary to switch brands because their favorite brand was no longer available. The vast majority of these switchers were customers of either Colgate or Proctor and Gamble, both of whom failed to provide a no-phosphate substitute carrying the same brand names as the popular phosphate products. Lever Brothers did provide such a substitute, and their customers constitute a major part of those who maintained brand continuity while switching to a no-phosphate product.

Respondents' estimates of the amount of detergent used per washload, the amount of extra ingredients used, and the cost of no-phosphate detergent were divided according to these "switcher" and "non-switcher" categories. The data are presented in Tables 9, 10, and 11. Table 9 indicates that nearly a third of those who had to switch brands felt the no-phosphate detergent required more per washload, while only four percent of the non-switchers thought so. Table 10 indicates a similar disparity regarding the amount of extra ingredients used. Table 11 indicates that over two thirds of the switchers believed no-phosphate detergent cost more to use while less than a third of the non-switchers felt this was the case.

It should be noted that in the case of both the switcher and the non-switcher, the consumer is using a totally new product, however, in one case she is also using a new brand, while in other she maintains brand continuity. The differential response according to brand continuity indicates the importance of the brand to the consumer, and in this case the loyalty appears to override the radical

changes in the physical and chemical contents of the product, even when this change is obvious to the consumer.

The results of this analysis tend to support the reactance theory because non-switchers perceived little limitation in their choice, while switchers were obviously limited. These differences in choice alternatives were directly manifested in evaluations of the amount of detergent used, the extra ingredients used, and the cost of no-phosphate detergent. In all three cases the differences are as predicted by reactance theory .

Table 9
Amount of Detergent Used^a
by Brand Continuity^b

Brand	N	More	Less	Same
Switched	51	32%	10%	58%
Not Switch	26	4%	28%	68%

Chi square = 9.41 df = 2 p < .01

^aDo you use more, about the same, or less (no phosphate detergent) for each washload than you did with (phosphate detergent)?

^bOf those respondents from Miami using no-phosphate detergent, some found it necessary to switch brands because their favorite brand was no longer available.

Table 10
Amount of Extra Ingredients Used^a
by Brand Continuity^b

Brand	N	More	Same or Less ^c
Switched	51	32%	64%
Not Switch	26	8%	92%

Chi square = 4.03 df = 1 p < .05

^aDo you use more, about the same, or less additional ingredients, such as bleach or fabric softener, with each washload when using (no-phosphate detergent) than you did with (phosphate detergent)?

^bOf those respondents from Miami using no-phosphate detergent, some found it necessary to switch brands because their favorite brand was no longer available.

^cCategories combined to provide sufficient expected value in cell.

Table 11

Cost of Non-phosphate Detergent^a
by Brand Continuity^b

Brand	N	More	Same or Less ^c
Switched	51	71%	29%
Not Switch	26	30%	70%

Chi square = 11.63 df = 1

p < .001

^aDo you think that no-phosphate (detergent) is more expensive to use, about the same, or less expensive to use than (phosphate detergent)?

^bOf those respondents from Miami using no-phosphate detergent, some found it necessary to switch brands because their favorite brand was no longer available.

^cCategories combined to provide sufficient expected value per cell.

If switchers are more reactive than the non-switchers, one would expect them to be more negative toward the law which limited their choices. Attitudes toward the law, by brand category, are presented in Table 12. While the majority of switchers favored the law, the proportion was significantly lower than the 91% of non-switchers who also indicated a favorable attitude. In this case, even though the data provide only marginal certainty of a significant difference, the results appear to be supportive of the theory of psychological reactance explanation of behavior by Miami respondents.

Table 12
 Attitude Toward the Law^a
 by Brand Continuity^b

Brand	N	Favorable	Unfavorable
Switched	45	69%	31%
Not Switch	23	91%	9%

Chi square = 3.10 df = 1 p < .10

^aHow do you feel about this recent law prohibiting the sale or use of products containing phosphates?

^bOf those respondents from Miami using no-phosphate detergent, some found it necessary to switch brands because their favorite brand was no longer available.

Summary

This report focused on differential patterns of response to questions concerning perception, attitudes, and evaluations, according to the respondent's area of residence and brand continuity. It appears that the theory of psychological reactance received general support from the data, and none of the results of the study directly contradicts the predictions of the theory. In view of the relative unimportance of laundry detergent, compared to other consumer products, the existence of signs of psychological reactance is indicative of the consumer's need to maintain her choice alternatives in the market. It seems likely that consumer reactance would be magnified directly with the increase in the number and kind of limitations imposed on consumer choice, and with the importance of the products subjected to restriction.

Footnotes

1. The research reported in this paper was supported by the College of Business Administration, University of Florida. The authors are indebted to Dennis C. Leslie for assistance in questionnaire design and in data processing.
2. Assistant Professor of Marketing, University of Florida and Assistant Professor of Marketing, San Diego State University, respectively.

References

1. Jack W. Brehm. A Theory of Psychological Reactance. New York: Academic Press, 1966)
2. Milton Rokeach. Beliefs, Attitudes and Values. San Francisco: Jossey - Bass, Inc., 1968, 156-178.
3. Julian B. Rotter. Generalized Expectancies for Internal Versus External Control of Reinforcement. Psychological Monographs, 1966, 80 (whole no. 609), 1-28.

ATTITUDES TOWARD WOMEN'S
LIBERATION AND PERCEPTION
OF ADVERTISEMENTS

Michael B. Mazis and Marilyn Beuttenmuller¹
University of Florida

In recent months the rhetoric of the women's liberation movement has permeated the mass media. One result has been a proposed constitutional amendment and a myriad of legislation dealing with women's rights.

Subtle discrimination against women has come into the spotlight. For example, the National Organization of Women (N. O. W.) contends that there are at least five ways that advertising, in any medium, subtly discriminates against women:

1. Women are shown as frivolous, docile, stupid, and incompetent.
2. Advertising appeals to women's insecurities and fears about their roles as wives, mothers, lovers, and house-cleaners. It encourages the image that the only role open to women is to cater to men. "Women are encouraged to have a manic obsession with cleanliness."
3. Advertising portrays limited and unrealistic stereotypes of male and female roles - both occupationally and sexually. This includes the sin of omission in showing only a limited number of occupational roles held by women. A recent survey of ads indicates that nine percent of ads showed women in working roles while ninety-one percent showed them in a decorative or homemaker roles. When shown in an occupation, most of the time it was as a secretary, nurse, or stewardess.
4. Women are used as decorative objects. Using women who serve no function except to decorate the product is dehumanizing, N. O. W. contends. When it is desirable to have a person in an ad, that person can be an active, competent woman, rather than a doll. They singled out industrial product advertising as a chief offender. Frequently, the industrial product is so uninteresting that the advertisers "throw in a good looking broad in some ridiculous pose."
5. Advertising tries to exploit the women's movement. Many ads are employing the word "liberation," implying association with the movement. Although it is basically a civil rights movement, it is subjected to much abuse "in the name of humor, creativity, and timeliness." (Chicago Tribune, 1971.)

In general, do women feel exploited by women's images in advertising or does the idea of exploitation even occur to women looking at advertisements? This paper attempts to provide a partial answer to the question. Specifically, this research attempt to determine whether a woman's attitude toward women's liberation will affect her perception of advertisements containing women.

The subject of attitudes and values affecting perception has been the focus of numerous psychological experiments. This emphasis on selective perception was termed the "new look" in perception by Bruner since it stresses "behavioral" influences on perception rather than emphasizing physical properties of stimuli (Bruner, 1958). While previous studies emphasized the influence of values (Haigh and Fiske, 1952; Postman, Bruner, & McGinnis, 1948; Solomon and Howes, 1950), income levels (Bruner and Goodman, 1947), skin color (Mark, 1943), and other factors on perception, the present study follows a pattern established by Kassarian (1965) in attempting to assess the impact of an internal psychological state on perception of advertisements.

Experiment 1

Methodology

For the experiment, five product categories were chosen - cleaning products, perfume, liquor, cigarettes and stereo equipment. Each category contained four advertisements taken from current magazines - one "unfavorable" ad, two neutral ads, and one "favorable" advertisement. Each of the five unfavorable ads embodies one of the aforementioned criticisms (although opposite favorable-unfavorable ads were not necessarily placed in the same product category). These twenty ads were so classified by a convenience sample of five female students at the University of Florida. The ads were disguised as to their actual brand names and any identifying labels, etc. Each product in a category was assigned a "brand name" of L, M, P, or H.²

A random sample of 24 female students were chosen from the 1971-72 University of Florida Student Directory. Data was collected by personal interview at the residence of each student. Each subject was given the ads for only one product category and she rated the women in the four ads on a series of nine six-point bipolar adjective scales chosen to be representative of the criticisms by N. O. W. (e. g., "competent-incompetent"). Also, the respondent indicated on a similar bipolar scale how willing she was to purchase the particular brand shown in the advertisements. After rating all four ads, each woman completed an attitude questionnaire consisting of 47 statements dealing with the major goals of women's liberation. The statements were scored on a five-point Likert-type continuum with items being scored in both directions to avoid acquiescence set. The attitude questions were administered after women viewed the advertisements in order to avoid biasing the ratings of the advertisements.

Results

The results shown in Table 1 indicate that coefficients for the women's liberation attitude test and the willingness to buy scores were .16 and .31 for the favorable and unfavorable ads respectively, both in the predicted direction, but failing to attain significance. Similarly, the correlation

coefficients for the women's liberation attitude test and the summative score of the nine bipolar adjective scales describing the woman in the picture were also in the predicted direction, .16 and .34, but not significant.

Table 1

Correlation of Women's Liberation Attitude Test
Willingness to Buy and Summation of Ten
Bipolar Adjective Scales

	<u>Willingness to Buy</u>		<u>Bipolar Adjective Scales</u>	
	Favorable Ads	Unfavorable Ads	Favorable Ads	Unfavorable Ads
Correlation Coefficient (r)	.16	.31	.16	.34
F-ratio	.49	2.31	.45	1.87

Experiment 2

Methodology

Realizing that the small sample size in experiment may have contributed to the lack of significant findings, a second study was conducted with a somewhat different methodology. Since large classes of students were used, advertisements were placed on slides. Slides were shown to two groups of approximately 50 and 100 female students. Both groups were shown the identical pictorial content in the ads; however, the copy **has changed** on four of the advertisements to reflect "favorable" and "unfavorable" statements towards women's roles in the ads. Four control ads were the same in both groups.

Respondents completed a six-point bipolar scale ("influential-not influential") for each of the eight ads and nine bipolar scales about the women pictured in each of the ads. Finally, subjects completed a revised 22 question women's liberation attitude questionnaire (appendix A). The attitude questionnaire was refined through the use of factor analysis, which revealed only one salient factor, discriminant analysis and examination of the intercorrelation matrix for the original instrument.

Results

In order to isolate extreme subjects, only respondents with the 10 highest and lowest attitude scores in group 1 and respondents with the 20 highest and lowest scores in group 2 were considered for analysis. These were subjects scoring in the upper and lower 20% of all women taking the women's liberation attitude test in their group.

Table 2
Influence Scores and Attitudes
toward Women's Liberation

Group 1

	Pro-Women's Liberation (10)	Anti-Women's Liberation (10)
Ad 1	1.4 ^b	2.0
Ad 2	3.1	3.9
Ad 3 ^a	4.0	4.6
Ad 4 ^a	4.3	3.8

^aPro-Women's Liberation

^bSix-point bipolar adjective scales with "1" for not influential and "6" for influential.

Table 2 shows that advertisements 1 and 2 were in the predicted direction with anti-women's liberation girls feeling that these two ads would influence them more than pro-women's liberation girls; however, the differences were not large enough to achieve statistical significance. Also, ad 4 was in the predicted direction, but ad 3 was in a direction opposite to that predicted. In general, these results are rather inconclusive.

For the second group, which contained twice as many subjects as the first group, the relationships are somewhat stronger, perhaps resulting from the larger sample size.

Table 3
Influence Scores and Attitudes
towards Women's Liberation

Group 2

	Pro-Women's Liberation (20)	Anti-Women's Liberation (20)
Ad 1 ^a	5.4 ^b	5.2
Ad 2 ^a	4.0	3.5*
Ad 3	2.3	2.8
Ad 4	2.0	2.4

^aPro-Women's Liberation Ad

^bSix-point bipolar adjective scales with "1" for not influential and "6" for influential

*p < .05

Influence ratings for all four ads in group 2 were in the predicted direction. Again, relationships are not very strong. Comparing the data across Tables 1 and 2, it is interesting to note that the pro-women's liberation ads were more influential than the anti-women's liberation ads. This may result because most girls attending college are rather career-oriented and even the anti-women's liberation group are more liberated than their counterparts who are secretaries or housewives.

Also of interest is the fact that married female college students showed some tendency to be more liberated than single students. While the number of married students was too small to reach any definite conclusions, this is a hypothesis worthy of further investigation.

Since the results from group 2 appear more stable, with a larger sample size and far smaller variance, subsequent analysis will deal with group 2 data only. Following the pattern established in the first experiment, the nine bipolar adjective scales describing the woman pictured in each advertisement were summed yielding one score per ad. In general, the same patterns were observed with the bipolar scales as with the influence scores - largely in the predicted direction but rather weak. Differences for ads 1 and 2 were significant at $p < .05$; however, due to the many t-tests used, these results should be viewed with caution.

Table 4

Adjective Scale Summation and
Attitudes Towards Women's Liberation

Group 2

	Pro-Women's Liberation (20)	Anti-Women's Liberation (20)
Ad 1 ^a	27.4	22.3*
Ad 2 ^a	32.6	28.4*
Ad 3	36.8	40.2
Ad 4	22.8	22.8

^aSix-point bipolar adjective scales with "1" for not influential and "6" for influential.

* $p < .05$

One of the interesting findings in this study is that pro-women's liberation girls have a more positive attitude towards advertising than the anti-women's liberation group. The results in Table 5 show that the four control ads were generally given higher scores by women with strong pro-woman's liberation attitudes.

Table 5
 Influence Scores and Attitudes
 towards Women's Liberation
 (control ads only)

Group 2

	Pro-Women's Liberation (20)	Anti-Women's Liberation (20)
Ad 5	5.1 ^a	4.2*
Ad 6	4.5	4.0
Ad 7	2.4	2.0
Ad 8	2.5	2.1

^aSix-point bipolar adjective scales with "1" for not influential and "6" for influential.

*p < .05

Conclusions

Overall, it appears that there is a weak relationship between the women's liberation attitude test and perception of the advertisements used in the two experiments. Considerable work remains to be completed in this area.

The women's liberation attitude test needs to be refined, with appropriate reliability and validity checks. External validity may be tested by administration to known groups of pro- and anti-women's liberation groups. The unidimensionality of attitudes towards women's liberation found in this study should be explored in subsequent studies.

A wider range of advertisements should be used. A non-college student population should be tested, particularly to measure the anti-women's liberation group. Finally, study is needed on the psychological reasons for women holding attitudes at both ends of liberation continuum. For example, it was found that the pro-women's liberation group consistently felt that the neutral ads would be more influential on them and they viewed the woman pictured in the ads more positively than the anti-women's liberation group. The underlying reasons for this phenomenon should be studied. These two experiments represent the start of trying to understand a complex area.

Footnotes

1. Michael B. Mazis is Assistant Professor of Marketing, University of Florida and Marilyn Beuttenmuller is a member of the Palm Beach County (Florida) Consumer Affairs Department.
2. These four middle-alphabet letters (consonants) were chosen because they have about the same frequency of use in English (stafford, 1966).

References

- Bruner, J. S. Social Psychology and Perception. In E. E. Maccoby, T. M. Newman and E. L. Hartley (Eds.), Readings in Social Psychology. New York: Holt, Rinehart & Winston, 1958, 85-94.
- _____ and Goodman, C. C. Value and Need as Organizing Factors in Perception. Journal of Abnormal and Social Psychology, 1947, 42, 33-44.
- Chicago Tribune, September 21, 1971, Section 3, 8.
- Haigh, G. V. and Fiske, D. W. Corroberation of Personal Values as Selective Factors in Perception. Journal of Abnormal and Social Psychology, 1952, 47, 394-98.
- Kassarjian, H. K. Social Character and Differential Preference for Mass Communication. Journal of Marketing Research, 1965, 2, 146-53.
- Mark, C. S. Skin Color Judgements of Negro College Students. Journal of Abnormal and Social Psychology, 1943, 38, 370-76.
- Postman, L., Bruner, J. S. and McGinnies, E. Personal Values and Selective Factors in Perception. Journal of Abnormal and Social Psychology, 1948, 43.
- Soloman, R. L. and Howes, D. H. Word Frequency, Personal Values and Visual Duration Thresholds. Psychology Review, 1951, 58, 256-70.
- Stafford, J. E. Effects of Group Influence on Consumer Brand Preferences. Journal of Marketing Research, 1966, 3, 68-75.

Appendix A

Revised 22 Question Women's Liberation Attitude Test

1. Women who have jobs demanding the same skills and responsibilities as a man's job but receive significantly less pay should accept the situation.
2. A man should be paid more for the same work than a woman if he is supporting a family and she is not.
3. Women are poor job risks emotionally and biologically.
4. Men are justified if they judge women on the basis of beauty and sex appeal.
5. Women are more likely to act illogically and "fall apart" in emergencies than men are.
6. Women should be prohibited from working overtime for their own good.
7. If a woman can play football as well as a man she should be on the team.
8. Toy trucks are suitable toys for both girls and boys.
9. A husband deserves more recreational time away from home than a wife does.
10. Housekeeping is woman's work.
11. Every woman should have equal job opportunity with men.
12. Every woman who wants one should be able to have an abortion.
13. State laws that prevent a woman from selling property or starting a business without her husband's consent are for the protection of women.
14. The sexual "double standard" as it now exists in our society is biologically determined.
15. Women are the weaker sex.
16. Every woman who wants it should have access to free or low cost child care centers.
17. Jobs that require hard physical labor should be open to women who are physically qualified.
18. Women should not be permitted to work at night.
19. Getting married is a social necessity for a woman.
20. Men who do household chores are unmasculine.
21. The use of "Mrs." and "Miss" are discriminatory for a women-"Ms." would be better.
22. Women should give up their jobs if their husbands want them to.

ARE WE MISSING THE MS.?¹

Beverlee B. Anderson²
University of Kansas

In 1963 Betty Friedan's book, The Feminine Mystique, was published. This marked the beginnings of the Women's Liberation Movement. Since that time much has been written about the Liberated Woman. Traditional women's magazines have featured articles on the working woman and new roles for women. Widely read magazines such as Life and Time have devoted entire issues to woman's changing role in society and her liberation from household duties.

As this movement has gained momentum various organizations have formed to spearhead some of the issues involved in the liberation. Membership in organizations such as NOW, WHEAL, Women's political Caucus and Women's Coalition has had tremendous growth. This indicates that more and more women are becoming concerned about the role of women in society. Many women are beginning to question the traditional roles they have been playing and the functions they have been performing.

These changing attitudes toward the role of the woman have presumably led to changing "self-concepts" of the women who have joined the liberation movement. Various researchers have studied the relationship between self-image and actions, including purchase behavior (Grubb and Hupp, 1968; Grubb, 1965; and Grubb and Grathwohl, 1967). The positive relationships between self-image and purchase behavior have led some researchers to claim that the self-concept is a meaningful mode of market segmentation.

If a woman who considers herself a liberated woman has a different self-concept than her non-liberated sisters, it is possible that her behavior patterns are different. These differences could then serve as a basis for market segmentation.

Methodology

To explore the possible behavioral differences between liberated and non-liberated women, personal interviews were conducted with 139 women in their homes. The respondents were selected from two sampling frames in Lawrence and Topeka, Kansas. One frame was a listing of all women employed by the State of Kansas in Topeka or the University of Kansas in Lawrence. One-half of the respondents were selected at random from this list. The remaining respondents were randomly selected from the telephone directories of Lawrence and Topeka, Kansas. Although the first sample frame is a subset of the second, no duplications occurred.

All respondents were between the ages of 18 and 75. Data was not collected on race; however, several of the respondents were considered to be minority group members. The average level of education was high school graduate, but the range was from sixth-grade education to the completion of a Ph.D.

The respondents were classed as "liberated," "not-liberated," or "undecided." This classifying was done by a question at the end of the interview in which each respondent was asked if she considered herself to be a liberated woman. The self-identification technique was used because it is believed that as a self-concept, there is no objective method which can be used to identify a woman as being liberated. This belief was supported by this research when it was found that liberated/non-liberated was not significantly related to 1) whether a woman works outside the home; 2) a woman's age; or 3) her level of

education. It was interesting to note that several older women, with an average level of education, who had never worked outside the home, considered themselves to be liberated. This is certainly in conflict with the stereotyped women's libber characterized by many articles.

Four areas of purchase behavior are explored in this research. These areas are 1) the shopping trip, 2) store preference patterns, 3) awareness or use of new products and 4) product information sources. These are areas where it was thought the liberated woman may be significantly different from other women.

The Shopping Trip

The trip to the grocery store has been a traditional role of the woman for many years. As role concepts change, it is possible that the trip to the grocery store is no longer the primary responsibility of the woman or perhaps this role is now shared with her spouse. To explore the behavior of the shopping trip, three aspects of the trip were examined, 1) the number of trips per week to the grocery store, 2) the number of grocery stores patronized on a regular basis, and 3) companions on shopping trips.

Trips per Week

If the liberated woman was less inclined to feel the trip to the grocery store was her responsibility, then it was thought that perhaps she would make fewer trips per week than would the non-liberated woman. Table 1 shows the average number of trips to the grocery store per week for the respondents. The figures show that most women tended to make only one trip per week to the grocery store. However, the chi-square statistic of 9.96 is significant at the .10 level. It is found that the liberated women tend to make more trips to the grocery store per week than was expected. Table 1 shows the observed and expected frequencies on the average number of trips per week.

Table 1

Average Number of Trips to Grocery Store per Week

	<u>1 Trip/Week</u>		<u>2 to 3 Trips/Week</u>		<u>3 or more Trips/Week</u>		<u>Total</u>
	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	
Liberated Women	40	48	34	27	4	3	78
Non-Liberated Women	20	18	9	10	0	1	29
Undecided	25	19	5	11	2	2	32
Total	85	85	48	48	6	6	139

Chi-square statistic, 4 degrees of freedom = 9.96
Significant at .10

These findings tend to refute the hypothesis that the liberated woman would make fewer trips to the grocery store per week than would other women. Since this question did not differentiate between major shopping trips and fill-in trips, it is possible that the liberated woman tends to make more fill-in trips, or it may be that she tends to buy in smaller quantities than other women, buying only what she needs rather than stocking her shelves.

Number of Grocery Stores Patronized

Some women like to do all of their grocery shopping at one store, while other women like to "shop around" at several stores in the area. It was thought that there may be a relationship between number of stores patronized and a woman's self image. Approximately 74 percent of the respondents shopped at only one store, while the remaining shopped at more than one. The chi-square statistic of .89 for 2 degrees of freedom is not significant, therefore, it is doubtful that there is any relationship between a woman's self image and the number of stores she patronizes. Table 2 shows this finding.

Table 2

Grocery Stores Patronized

	Only One Store	More Than One Store	Total
Liberated Women	58	20	78
Non-Liberated Women	23	6	29
Undecided	22	10	32
Total	103	36	139

Chi-square statistic with 2 degrees of freedom = .89
Not significant at .10

Companions on Shopping Trips

Some women are independent and like to go shopping alone, while other women prefer to have companions on shopping trips. Since many liberated women are thought of as being more independent than other women, it was thought that the self-identified liberated women may tend to shop alone.

About half of the respondents (67) went shopping alone, while 72 of the respondents had companions on their shopping trips. When comparing the companions of liberated and other women, it is found that the liberated women are just as likely to have companions on their shopping trips as are the non-liberated or undecided women.

Of the 72 women who said they did have companions on shopping trips, it was found that the majority of these women went with their husbands or children. Very few of the respondents went shopping with friends or relatives. However, there does not appear to be any relationship as to a woman's liberated status and her choice of companions on a shopping trip. Table 3 shows the frequencies of various companions on shopping trips.

The findings indicate that when examining shopping trip behavior, the number of trips per week to the grocery store is the only area where we find a significant difference between the self-identified liberated women and other women. The other two aspects of the shopping trip analyzed, number of stores patronized and companions on shopping trips, do not show significant differences between the liberated and other women.

Table 3
Companions on Shopping Trips

	Friends	Husband	Children	Relatives	Total
Liberated Women	4	12	15	2	33
Non-Liberated Women	1	8	8	0	17
Undecided	1	11	8	2	22
Total	6	31	31	4	72

Chi-square statistic, 6 degrees of freedom = 3.51
Not significant at .10

Store Preference Patterns

There are many evaluative criteria a woman may use in choosing the grocery store she will patronize. The relative importance of seven criteria were examined in relation to store preference. The seven criteria evaluated were: 1) location, 2) prices, 3) large assortment, 4) store personnel, 5) quality of meat and/or produce, 6) efficiency and speed of shopping, and 7) store decor. These seven items were rated on a 1-6 Likert Scale with a 1 being extremely important and a 6 being extremely unimportant.

Of the seven factors examined, prices and quality of meat/produce were the most important criteria in store patronage while store personnel and store decor were the least important.

Table 4 shows the relative importance of each of the factors and the chi-square statistic when the ratings of the three liberation classifications are compared. The only factors which had significant chi-square statistics were location and speed and efficiency. Tables 5 and 6 show the responses of the various classifications of women on these factors.

Location as a factor is more important than expected to the "undecided women," while it is less important than expected to the non-liberated women. Speed and efficiency also appear to be more important to the liberated women. When considered together, these two factors indicate that the non-liberated women are probably less concerned about convenience and speed of shopping than are other women. This could be explained by the possibility that the non-liberated woman considers the trip to the grocery store to be an important part of her role; therefore, she is not so concerned about speed and convenience.

Table 4

Importance of Evaluative Criteria in Store Preference

Order of Importance	Factor	Chi-Square Statistic
1	Price	2.78
2	Quality of meat and/or produce	3.77
3	Assortment	6.41
4	Location	25.67 ^a
5	Efficiency and Speed	19.17 ^a
6	Store Personnel	14.10
7	Store Decor	10.99

^aChi-square significant at .10

Table 5

Location as an Evaluative Criteria in Store Preference

	Extremely Important			Extremely Unimportant			Total
	1	2	3	4	5	6	
Liberated Women	37	26	11	3	0	1	78
Non-Liberated Women	4	15	3	5	1	1	29
Undecided	21	8	2	1	0	0	32
Total	62	49	16	9	1	2	139

Chi-square statistic, 10 degrees of freedom = 25.67
Significant at .10

New Products

On the assumption that liberated women may be more "venturesome" (Rogers, 1962, p. 189) than other women, it was thought that they may be more likely to have heard of and purchased new products. To examine this possibility the respondents were asked if they had heard of or purchased seven new products which are carried in grocery stores. At the time of this survey, all products

Table 6
Speed and Efficiency of Service as an Evaluative
Criteria in Store Preference

	Extremely Important			Extremely Unimportant			Total
	1	2	3	4	5	6	
Liberated Women	29	32	13	2	2	0	78
Non-Liberated Women	7	8	7	2	4	1	29
Undecided	10	15	2	4	1	0	32
Total	46	55	22	8	7	1	139

Chi-square statistic, 10 degrees of freedom = 19.17
Significant at .10

had been on the market for less than four weeks. The seven new products tested were: Max Pax, Aires Tissues, Armour Taco Meat Spread, Stretch and Seal, Campbell's Chunky Soup, Soft Plus Fabric Conditioner, and Corn Diggers. It was thought that these products covered a range of products normally purchased at the grocery store; therefore, were representative of new grocery products.

Aggregating all seven new products it is found that there is no significant difference between liberated and other women in their awareness of or purchase of new products. Table 7 illustrates these findings.

Table 7
Awareness/Purchase of New Products

	Unaware of Product	Heard of Product	Purchased	Total
Liberated Women	276	222	48	546
Non-Liberated Women	96	85	22	203
Undecided	106	92	26	224
Total	478	399	96	973

Chi-square statistic, 4 degrees of freedom = 2.13
Not significant at .10

Apparently the liberated women are no more likely to have heard of or be innovators with new grocery products than are women in general.

When each of the new products was analyzed individually, Max Pax and Chunky Soup had significant chi-square statistics. With Max Pax, fewer than expected non-liberated women had heard of the product, while more than expected "undecided" women had heard of the product. When analyzing the Soup finding, a different pattern emerges. With this product fewer than expected undecided women had purchased the soup, while more than expected non-liberated women had purchased the soup. Both of these differences concern the non-liberated and the "undecided" women, rather than the liberated women. So, regardless of the products tested, the liberated women did not tend to be more likely to have heard of or purchased these products than other women.

Product Information Sources

There are many sources of information available to the consumer today. Research has shown that not necessarily the same sources are used for finding out about new products and getting a consumer interested in trying a new product (Robertson, 1971, p. 155). For this reason, two aspects of information sources were examined: those which are most important in finding out about a new item on the market and the importance of information sources in getting a person interested in trying a new item for the home. Six information sources were analyzed in this research, magazine advertising, friends and relatives, co-workers, television advertising, newspaper advertising and neighbors.

Finding Out About A New Product

Table 8 shows the relative importance of each information source in finding out about a new product. Friends and relatives and TV advertising were ranked very closely in relative importance by the respondents. One possible explanation why co-workers ranked so low is that not all of the respondents worked outside the home; therefore, many of the respondents did not have any co-workers with whom they could identify.

The chi-square statistics in Table 8 indicate the statistics which were found when comparing the importance of each information source among the three liberated classifications. Magazine advertising was the only information source where there was a significant difference between the three groupings of women. Table 9 shows these frequency distributions.

The analysis of Table 9 shows that liberated women tend to consider magazine advertising less important than expected while the undecided women tended to rate it more important than expected. The analysis of TV advertising, although not significant, tends toward similar findings. It appears that the liberated woman does not consider TV and Magazine Advertising, two of the most important marketing dominated information sources, as important as women who do not consider themselves to be liberated. It may be that she is a little less attentive to advertising messages, or she may tend to give less credence to them.

Interested in Trying a New Item

The relative importance of each of the six information sources in getting a person interested in trying a new item are shown in Table 10. Friends and relatives and TV advertising are rated as the most important information sources in this analysis. Magazine and newspaper advertising are considered to be the least important sources.

Table 8
Information Sources for Finding Out About New
Products on the Market

Rank in Importance	Information Source	Chi-Square Statistic
1	Friends and Relatives	14.95
2	TV Advertising	13.30
3	Neighbors	6.23
4	Magazine Advertising	17.67 ^a
5	Newspaper Advertising	7.81
6	Co-workers	10.93

^aChi-square significant at .10 level

Table 9
Importance of Magazine Advertising in Finding
Out About New Products

	Extremely Important			Extremely Unimportant			Total
	1	2	3	4	5	6	
Liberated Women	9	25	22	13	4	5	78
Non-Liberated Women	2	6	9	4	7	1	29
Undecided	4	14	11	2	1	0	32
Total	15	45	42	19	12	6	139

Chi-square statistic, 10 degrees of freedom = 17.67
Significant at .10

The significant differences between groups of women as related to friends and relatives and TV advertising are attributable to differences between the non-liberated and undecided groupings of women. In both instances, the undecided women tended to rate these information sources less important than expected, while the non-liberated women rated these sources more important than expected.

Even though there were some significant differences among the groups and evaluations of the importance of information sources, it cannot be said that

Table 10
Information Sources for Getting Interested
in Trying a New Item

Rank in Importance	Information Source	Chi-Square Statistic
1	Friends and Relatives	18.28 ^a
2	TV Advertising	16.09 ^a
3	Neighbors	15.82
4	Co-workers	11.37
5	Newspaper Advertising	11.71
6	Magazine Advertising	14.24

^aSignificant at .10 level

liberated women tend to use different information sources than do other women. So from a practical standpoint of market segmentation, it seems unlikely that we could segment these women on the basis of the information sources they use.

Implications for Future Research

In this research self identification as a liberated or non-liberated woman was used as a surrogate indicator for self image in the hopes that a relationship would exist between self-image and purchase behavior. This information could then possibly be used as a basis for market segmentation. The results of this exploratory research were not clear cut. It appears that there are some differences between the liberated woman and other women, such as the number of grocery trips per week and importance of some factors in choosing the store they will patronize. However, the reasons for these differences are not explained by this research. As exploratory research, this has just scratched the surface of some possible differences among women.

Liberation is not an absolute, but a continuum; therefore, by treating liberation as an absolute, it is possible that the differences in the extreme positions on liberation are clouded by the more moderate women. In future research, a scale or continuum would probably be a better instrument for measuring liberation, if a unidimensional measure is used.

It is also possible that a multidimensional measure of views on women's liberation would give a more accurate picture of a woman's self image. Future research in this area would probably be more fruitful if a multidimensional measure was used and if attitudes toward shopping as well as behavior were studied.

Footnotes

1. This research was supported in part by the School of Business Research Fund, University of Kansas.
2. Beverlee B. Anderson is Assistant Professor of Business Administration, The University of Kansas.

References

- Grubb, E. L. Consumer perception of 'self concept' and its relationship to brand choice of selected product types. Proceedings, Winter Conference, American Marketing Association, 1965, 419-422.
- Grubb, E. L. & Grathwohl, H. L. Consumer self-concept, symbolism and market behavior: a theoretical approach. Journal of Marketing, 1967, 31, 22-27.
- Grubb, E. L. & Hupp, G. Perception of self, generalized stereotypes and brand selection. Journal of Marketing Research, Feb., 1968, 58.
- Rogers, E. M. Diffusion of innovations. New York: The Free Press, 1962, 189.
- Robertson, T. S. Innovative behavior and communication. New York: Holt, Rinehart and Winston, Inc., 1971, 155.

THE FEMINIST MOVEMENT; ATTITUDES BEHAVIOR AND POTENTIAL

Alice T. Beery

Introduction

The feminist movement has experienced rapid growth and increased visibility during the last few years. Existing groups such as the National Organization of Women (NOW) have grown and new groups such as the National Women's Political Causes have been formed. Organized women played an active role in both the Republican and Democratic conventions.

The feminist movement has had several obvious manifestations. Media exposure of the movement has not only increased but it has also changed in nature. Articles about the feminist movement frequently appear in a broad spectrum of newspapers and magazines. Initially such articles were covered as special or feature news, however, they are now covered as straight news. Even more significant is the establishment of a national feminist magazine, Ms. The effect of the feminists is also reflected by the fact that the Equal Rights Amendment has been ratified by 21 state legislatures, and is before many more.

While the media reflects the views of national spokeswomen such as Germaine Greer, Betty Freidan, and Gloria Steinem, attitudes and opinions of women in general toward the feminist movement are not easily assessed. The purpose of this study is twofold: 1) to investigate women's view of the feminist movements in general, and especially the equal rights issue, and 2) to examine strategies for influencing women to take a more active part in determining their role.

To achieve these objectives an attempt is made to answer the following questions:

1. Do women perceive discrimination to be a problem?
2. What explanations can be offered for women's perceptions of discrimination?
3. What strategies of influence are suggested by the analysis?

Ninety-four women chosen at random from the Columbus, Ohio telephone directory participated in a twenty minute telephone survey. They were asked questions regarding their attitudes toward discrimination, feminist groups, the Equal Rights Amendment, sex roles, family behavior patterns, self-confidence and demographic characteristics. Forced choice, sentence completion, and Liekert-type questions were included. As the small sample indicates, this investigation is a pilot study, and any conclusions are necessarily tentative.

Perceptions of Sex Discrimination

Two types of sex discrimination were investigated in this study; 1) job related discrimination and 2) sex role discrimination. Job related discrimination is quite overt. It includes discriminatory hiring practices, unequal pay scales for women and men, and discrimination in opportunities for promotion. On the other hand, sex-role discrimination takes much subtler forms. It includes such things as the frequent assumption by men that women are only qualified to do menial tasks such as housework, or that women's role is that of cooking and cleaning for men.

Job Discrimination

Description - The average respondent perceived discrimination in job related areas (Table I, Questions 1-4). She felt that men should not receive higher pay than women for the same work merely because they have a family to support. While she was not convinced that employers discriminate in hiring practices (36.2% felt they did, 38.3% felt they did not), she did feel that once on the job women did not receive the same pay for the same work, and did not have the same opportunities for advancement.

Explanation - It can be hypothesized that a woman is likely to perceive job discrimination because of its overt nature. She is likely to have witnessed or experienced such discrimination.

Sex Role Discrimination

Description - Sex role discrimination, however, is much more difficult to pinpoint. It is the feeling of inferiority that comes from being delegated menial tasks such as housework and it is the result of a lifetime of socialization. Women exhibited ambivalence about sex role discrimination. Although they indicated a desire to break out of traditional female sex roles they still believed they should perform many of the home related tasks.

Several questions indicated a desire to break away from the traditional female sex role. The women gave indications of a desire for independence for themselves and for their daughters. In response to the sentence completion question, "The thing I regret most . . .", 21.7% regretted getting married so young and 26.0% regretted not finishing college. In response to the question, "From a very young age girls should be taught to . . .", 43.0% stated girls should be taught self-reliance and independence.

Women also indicated a desire for a more equalitarian role. When asked to complete the question "If I could convince men of one thing it would be . . .", 45.1% responded with one of the following--equal rights for women, women are as intelligent as men, women have their own life to live, and men should make marriage more of a partnership.

The women surveyed also indicated strong support for issues which would free them from family responsibility--85.1% supported family planning, 88.8% accepted the pill as a method of birth control, 60.7% indicated support for women's right to abort, and 76.6% felt day care centers provided adequate care.

In spite of these indicators of the desire for sex role change, most respondents identified strongly with the traditional feminine role. The majority of women felt their most important role was that of either mother or housewife (Table I, Question 5). Only 4.3% stated self-fulfillment as their most important role. Consistent with the belief that the mother role was most important was the opinion that women with young children should not work (Table I, Question 6). Women's attitudes regarding family roles are consistent with their behavior. This is indicated by the significant association between various roles and actual behaviors. The relationship between who should perform traditional female roles such as washing dishes and caring for children, and who does, in fact, perform these roles was significant at the .01 level. Women also accepted the traditional value

that women should be married by a given age. 50.0% answered the question, "Women should get married when . . .", with a specific age between 20 and 26 years.

Explanation - Why did the same women who recognized job discrimination fail to recognize sex role discrimination? Three factors may explain this phenomena. First, women's attitudes may have been strongly influenced by what they perceived their husbands' attitudes to be, (70.2% of the respondents were married). Second, the women may have been comfortable with the traditional female sex role. Third, women may lack exposure to change agents and change strategies which communicate the existence of sex role discrimination.

Husband Domination

Women perceive men (their husbands) as rejecting women's groups. In response to the question, "Men feel that women's groups are . . .", 83.3% stated that men feel women's group are "silly" or "a bunch of gossips." In addition, the average respondent felt that her husband or family would not encourage her to join a feminist group, although they might verbally support equal rights for women (Table I, Questions 7, 8). They also felt that their husband or family would not help out so that they could participate in feminist group activities (Table I, Question 9). Several women volunteered that the reason they would not join a feminist group was that their husband would not permit it.

It is reasonable to assume that married women are influenced by their husbands. This study indicates that the majority of women perceive their husbands as viewing the feminist movement negatively. Balance theory states that if a wife loves her husband (+) and the husband hates the feminist movement (-) then in order to maintain stability in the relationship the wife must also hate the movement (-). (Figure I). This was supported by the finding that divorced or separated women (women not influenced by husbands) were more likely to support the feminist movement.¹ While this admittedly oversimplifies the relationship it does indicate the importance of the husband.

Self Confidence

The women indicated a high degree of self confidence as measured by the questionnaire. (Table II). This confidence may reflect the fact that women feel competent in the traditional female sex role. The role may be boring, menial, and degrading, but they are able to perform it skillfully and therefore have a high degree of self-confidence. Changing their role threatens their self-confidence and consequently they resist change.

Lack of Exposure

Even those women who are potential supporters of the feminist movement may remain apathetic due to selective exposure. The selective exposure hypothesis states that individuals are most frequently exposed to information supportive of their initial position. Moreover, selective exposure does not result from active avoidance of conflicting information, but rather from the fact that the individual most frequently gravitates to an environment where he receives supportive information. To avoid selective exposure,

information must appear in media to which the target audience is exposed. The study indicates that women are not exposed to information regarding the feminist movement. When asked to name any feminist groups they could think of, 83.0% of the respondents could not name any. Even aided recall of names of feminist groups was low (Table III). Although 61.7% of the women were active in some organization, none were active in a feminist organization. Many stated they would not join a feminist group because they were uninterested, they questioned the ideas, they lacked time, or they were happy with their current situation. However, the women did not express a strong negative attitude toward the movement. They agreed that the ideas of the feminist groups were good but they did not feel they belonged in the movement (Table I, Question 10). In general they lacked the conviction and commitment needed to join a feminist group.

The women also lacked knowledge about the Equal Rights Amendment, (ERA). Only 61.7% had heard of the ERA and of these 34.5% strongly supported it while the majority (58.6%) did not feel strongly one way or another. For those women exposed to information about the ERA, comprehension was low (Table IV). The majority of respondents could not answer questions about the effects of the ERA. However supporters of the amendment did show higher comprehension than non-supporters.

It appears that the low level of exposure to information about feminist groups and issues may be a major factor in women's low perception of sex role discrimination. Information about the movement is frequently found in feminist publications (Ms, Now Newsletter, Spokeswoman), while the average respondent stated she reads traditional female magazines (Ladies Home Journal, McCalls, Family Circle).

Summary

Women perceive overt discrimination such as job discrimination, however are not likely to perceive the more subtle sex role discrimination. There are several explanations for this finding. Married women may be influenced by what they perceive as negative attitudes of their husbands toward feminist groups advocating an equalitarian status for women. Women may also be comfortable with the traditional sex role. They are confident of their ability to perform the traditional role and change threatens their confidence. Finally, women may lack information about feminist groups. It is likely that they are selectively exposed to information supporting the traditional female role and are not exposed to information about new roles for women. In spite of their adherence to traditional roles, women exhibit some latent desires for independence. Therefore, it is likely that these women could be persuaded to take a more active part in determining their role. Moreover, their daughters are even more likely to define a new role for themselves. They are exposed to a conflict between the role their mother teaches them (independence, self-reliance) and one she actually plays (traditional female role of submission). Tomorrow's generation will be forced to resolve the conflict and define a new female role.

Strategies for Change

Given this profile of women, several strategies for change can be suggested. Strategies for change may be dichotomized into media vehicle and message content strategies.

Media Vehicle Strategies

As was noted earlier, the typical woman has not had extensive exposure to information about the feminist movement and issues. Mere exposure may increase acceptance, such as in the case of the Equal Rights Amendment. Therefore, to increase acceptance one must increase exposure. To increase exposure, articles dealing with feminist issues should appear in sources such as traditional feminine magazines (Ladies Homes Journal, Family Circle, McCalls) to which these women are known to be exposed.

Explanation of the feminist movement and issues involves complex arguments. When complex arguments are necessary a print media is recommended. Such media increases comprehension by enabling the reader to digest the material at his own speed.

Finally, the credibility of the source may also moderate persuasion. Magazines with which the woman is familiar and trusts should be used. Magazines such as Ladies Home Journal, McCalls and Family Circle are likely to be trusted by the change target. With repeated presentation it is expected that the positive affect of these media is increasingly transferred to the communication.

Message Content Strategies

The elements of the appeal such as its source and content also moderate acceptance. Advertisers frequently seek to induce persuasion through use of source enhancement strategies such as using an expert or trustworthy source. National spokeswomen for the feminist movement have been individuals such as Betty Freidan and Gloria Steinem. The credibility of such sources with the relevant audience is questionable and should be tested. Sources which have high credibility should be used. Moreover, sources may increase credibility through co-orientation. Co-orientation may be achieved by first developing arguments on an issue supported by the audience (such as family planning). The trustworthiness so induced may then be effective in persuading women about issues on which they have negative or neutral viewpoints (such as the ERA).

The attitudes of the husband should be investigated. The wife's perception that the husband views the feminist movement negatively may be a self-fulfilling prophecy. The woman's perception of the husband's negative attitude may foster such an attitude toward increased freedom and actually reduce her independence. If this is true, then, perhaps messages and programs directed at consciousness-raising for both husband and wife would be most appropriate. If the husband's negative attitude is real rather than merely perceived strategies to change his behavior are needed. Consistent with dissonance theory, attempts to change behavior may precede attitude change. Similarly strategies to change the wife's behavior and thus attitude may be employed.

Footnote

1. The sample of divorced and separated women was too small to draw any firm conclusions.

Table I
Responses to Survey Questions

Question 1: Because men often have a family to support, they should receive higher pay than women for the same work.

Agree: 27.7%
Disagree: 68.1%
Neither: 4.3%

Question 2: In hiring, employers discriminate against women.

Agree: 36.2%
Disagree: 38.3%
Neither: 25.5%

Question 3: Women generally receive the same pay as men when they perform the same work.

Agree: 21.3%
Disagree: 63.8%
Neither: 14.9%

Question 4: Women generally have the same opportunity for advancement as men.

Agree: 27.7%
Disagree: 57.4%
Neither: 14.9%

Question 5: What is your most important role as a woman?

Mother: 38.3%
Housewife: 29.8%
Wife: 10.6%
Self-fulfillment: 4.3%
Other: 17.0%

Question 6: Women with young children should not work.

Agree: 46.8%
Disagree: 27.7%
Neither: 25.5%

Question 7: The man or family I know best believes in equal employment for women.

Agree: 57.4%
Disagree: 29.8%
Neither: 12.8%

Question 8: The man or family I know best would encourage me to join a feminist's group.

Agree: 17.0%
Disagree: 68.1%
Neither: 14.9%

Question 9: The man or family I know best would help out so that I could participate in feminist's group activities.

Agree: 34.0%
Disagree: 51.1%
Neither: 14.9%

Question 10: I feel that feminist's groups generally have good ideas but I don't feel I belong in the movement.

Agree: 53.2%
Disagree: 23.4%
Neither: 23.4%

Table II

SELF-CONFIDENCE QUESTIONS

Per cent

	HOW OFTEN DO YOU FEEL YOU ARE TO BLAME FOR YOUR MISTAKES? IS IT
17.0	VERY OFTEN
27.7	FAIRLY OFTEN
34.0	SOMETIMES
8.5	ALMOST NEVER
	HOW OFTEN DO YOU HAVE THE FEELING THERE IS <u>NOTHING</u> YOU CAN DO WELL? IS IT
4.3	VERY OFTEN
4.3	FAIRLY OFTEN
46.8	SOMETIMES
31.9	ALMOST NEVER
	HOW OFTEN DO YOU FEEL YOU DISLIKE YOURSELF? IS IT
6.4	VERY OFTEN
6.4	FAIRLY OFTEN
34.0	SOMETIMES
42.6	ALMOST NEVER
	WHEN IN A GROUP OF PEOPLE, DO YOU HAVE TROUBLE THINKING OF THE RIGHT THINGS TO TALK ABOUT? IS IT
4.3	VERY OFTEN
12.8	FAIRLY OFTEN
36.2	SOMETIMES
36.2	ALMOST NEVER
	DO YOU EVER FEEL SO DISCOURAGED WITH YOURSELF THAT YOU WONDER WHETHER ANYTHING IS WORTHWHILE? IS IT
6.4	VERY OFTEN
2.1	FAIRLY OFTEN
40.4	SOMETIMES
10.6	ALMOST NEVER
	DO YOU EVER FEEL AFRAID WHEN YOU ARE GOING INTO A ROOM BY YOURSELF WHERE OTHER PEOPLE HAVE ALREADY GATHERED AND ARE TALKING? IS IT
2.1	VERY OFTEN
6.4	FAIRLY OFTEN
27.7	SOMETIMES
55.3	ALMOST NEVER

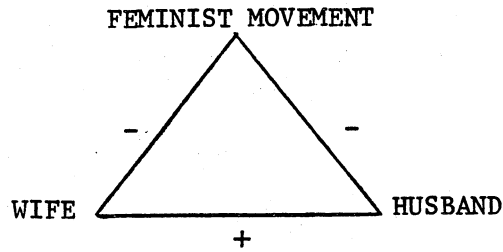
Table III

AIDED RECALL OF FEMINIST GROUPS

HAVE YOU HEARD OF:	YES FAVORABLE	NO UNFAVORABLE	NO
NATIONAL ORGANIZATION OF WOMEN	14.9	6.4	76.6
WOMEN'S EQUITY ACTION LEAGUE	4.3	2.1	91.5
LEAGUE OF WOMEN VOTERS	78.7	2.1	14.9
OHIO COMMISSION ON THE STATUS OF WOMEN	6.4	4.3	87.2
W.I.T.C.H.	4.3	2.1	91.5

Figure 1

WOMEN ARE INFLUENCED BY THEIR HUSBANDS



BALANCE THEORY

Table IV

KNOWLEDGE ABOUT THE ERA

UNDER THE ERA, WOMEN WOULD NOT BE GRANTED ALIMONY. THIS STATEMENT IS

TRUE 14.3% *FALSE 14.3% DON'T KNOW 71.4%

UNDER THE ERA WOMEN WOULD NOT GET CUSTODY OF THEIR CHILDREN. THIS IS

TRUE 3.6% *FALSE 39.3% DON'T KNOW 57.17%

SPECIAL RESTRICTION ON PROPERTY RIGHTS OF MARRIED WOMEN WOULD BE UNCONSTITUTIONAL UNDER THE ERA. THIS STATEMENT IS

*TRUE 17.9% FALSE 3.6% DON'T KNOW 78.6%

THE ERA WOULD ELIMINATE SEPARATE RESTROOM FACILITIES IN PUBLIC BUILDING. THIS STATEMENT IS

TRUE 14.3% *FALSE 25.0% DON'T KNOW 60.7%

WOMEN WOULD BE EQUALLY SUBJECT TO THE DRAFT AS MEN UNDER THE ERA. THIS STATEMENT IS

*TRUE 21.4% FALSE 7.1% DON'T KNOW 71.4%

WOMEN WOULD BE EQUALLY SUBJECT TO JURY DUTY AS MEN - UNDER THE ERA. THIS STATEMENT IS

*TRUE 50.0% FALSE _____ DON'T KNOW 50.0%

WORK LAWS THAT APPLY ONLY TO WOMEN WOULD BE UNCONSTITUTIONAL UNDER THE ERA. THIS STATEMENT IS

*TRUE 21.4% FALSE 26.7% DON'T KNOW 67.9%

UNDER THE ERA WOMEN WOULD BE REQUIRED TO LIFT HEAVY WEIGHTS. THIS STATEMENT IS

TRUE 17.9% *FALSE 25.0% DON'T KNOW 57.1%

*Indicates the correct response.

AN EMPIRICAL INVESTIGATION OF THE VALIDITY OF SELECTED
ATTITUDE AND ACTIVITY MEASURES

Albert V. Bruno¹ and Edgar A. Pessemier²

All investigators, particularly those working in the social sciences, should be concerned about measurement problems. Prior to data collection, judgment must be exercised in the design or selection of instruments. After data has been collected, the quality of the data must be appraised. In exercising prior judgments and examining the quality of collected data, the researcher is interested in reliability (accurate, consistent data) and validity (data that measures what it is designed to measure).

Historically, marketing researchers have not demonstrated great concern about the reliability or validity of the instruments which they have employed. Cost considerations, time limitations or the nature of the applied work usually have prevented practitioners from adopting the traditions of careful measurement characteristics of basic social science research. On the other hand, academic investigators in marketing and consumer behavior have been scandalously short of the resources required to investigate the range of important topics related to reliability and validity. To fill a small part of this void, we recently reported on the reliability of a series of studies conducted by researchers affiliated with Purdue University (Bruno & Pessemier, 1971).

In this validity study, we examine a similar body of research. These studies measure many important life-style, "psychographic" components of consumers' activities and attitudes. Naturally, the decision to use studies generated largely by the "Purdue group" of investigators was motivated by the local availability of raw data, source documents and useful unpublished results.

The Validity of Measurements

The validity of measurements made by an instrument depends on the scientific or applied purpose(s) of the user. One researcher may be interested in testing a descriptive theory about the common properties of classes of consumers or enterprises; another researcher may be interested in measurements which predict selected types of market behavior. In each case, validity must be judged against some appropriate criterion. The criterion and associated methods for judging validity can be conveniently classified into four important types:

Construct validity: This form of validity must be examined in light of adequacy and the results of three steps: the extent to which

- (1) theoretically satisfying constructs can account for test performance
- (2) hypotheses have been developed that flow from the theory or construct(s) and
- (3) the hypotheses are subjected to empirical test.

Content validity: The sampling adequacy or degree to which the test is representative of the construct or properties being measured.

Concurrent validity: "Immediate predictive" validity indicating the common variability of the test outcomes (measurements) and the one or more criterion variables. The passage of time and/or the order of causality are not necessary elements.

Example: One section of an instrument may be designed to measure product-specific innovativeness and another section may contain a previously employed test of product-specific information seeking and transmitting. Since prior theoretical considerations provide insights about the expected common variability of these two tests, concurrent validity of the new

measure of product-specific innovativeness can be examined in part by the correlation of the test outcomes.

Predictive validity: The common variability of the test outcomes (measurements) and one or more criterion variables when the passage of time and/or order of causality are necessarily considered.

Example: Individuals' brand preferences are scaled to predict their subsequent choice of brands. The measurement of brand preferences and the observation of brand choices are made at different times and in a natural time order. Note, however, that the time order would be reversed if one were trying to predict preference from choice. Finally, if changes in preference resulting from brand choice and/or use were of interest, then two preference measures and one choice or use experience would be required. In all these cases, however, the researcher is concerned with predictive validity.

In this paper we will examine essentially one type of validity, concurrent validity. Our reasons for doing so are practical. The discussion of the construct validity of each set of measurements would represent an impossible task, one in which each researcher has a better chance of success individually than we could hope to achieve in summary form. The interested reader is encouraged to consult the primary sources and to judge construct validity for himself. Because the nature of the constructs is not easily summarized, content validity must also be by-passed in this paper. Since the exploratory context of these studies reduces interest in the problem of content validity, the loss is small.

Predictive validity is not seriously treated in this discussion because of inadequate data. In some cases, investigators employ self-reported prior behavior or expected future behavior. These cross-sectional measures do not constitute "hard" tests of predictive validity, particularly for the classes of measures which are examined here. Therefore, we will focus on concurrent validity, accepting the exploratory nature of much of the research and looking across broad sets of criterion variables of potential interest to marketing scholars and managers.

Data Bases

Recently, an evolving area of data collection and research activity has been concerned with market-related measures of consumers' activities and attitudes. Although the evolutionary time has been short, a large number of related life-style, psychographic studies have appeared. The sheer volume of this work makes it impractical to summarize results within the confines of this discussion. Interested readers should consult recent articles by Hustad and Pessemier (Hustad & Pessemier, 1971) and Kassarian (Kassarjian, 1971) for a critical review of the subject.

At this point in the evolutionary process, it is not clear whether activity and attitude measures (more narrowly described as life-style or psychographic variables) will mature into well-developed, widely used bases for consumer and market analysis or will play a limited, specialized role. What is clear, however, is the need for more extensive validity and reliability research in order to allow this emerging area to develop its natural theoretical and applied stature. We believe that Kerlinger's remark ". . . all psychological and educational measuring instruments must be critically and empirically examined for their reliability and stability" applies equally to marketing and consumer research and the validity of the instruments employed in these fields (Kerlinger, 1965).

Specific Studies and Measurements

Although concurrent validity in a marketing setting concerns practical issues and outcomes, difficult problems in validation are associated with the selection of criteria. For example, if a marketing researcher is interested in enlarging the psychological profile of potential triers of a new product, some dimension of new product purchasing behavior must be employed as the criterion. However, if a marketing researcher is interested in developing measures which are associated with determining the susceptibility of consumers to persuasive communication, he may find measuring latent persuadability for a class of communications is an imposing task. In this case, the validity of the criterion variable(s) may be doubtful.

In the set of activity and attitude studies discussed in this paper, we examined criteria which are reported market behavior or widely accepted attitudinal indices such as brand preferences. The variables tested against these criteria are more exploratory in nature; they are drawn largely from emerging specialized segments of the personality, life style and value domains. The reliability of selected sets of these activity and attitude measures have been examined. The wider, overlapping set of studies used as data sources for this paper are summarized in Table 1. Since the investigators who conducted the investigations have described their data collection methods, these details will not be discussed here.

Classifying the measures that appear in the last column of Table 1 is a complex, elusive problem. The taxonomic scheme shown in Table 2 is adapted from DeBruicker, 1971. The first row is concerned principally with measures of properties that are highly enduring such as personality and demographic characteristics, environmental conditions and some ownership profiles. The second and third rows are concerned with consumer predispositions to action subject to influence and change. These attitudes are separated into perceptual (cognitive) and preference (affective) elements. Finally, the fourth row is concerned solely with manifest behavior. The columns are arrayed along a continuum of properties ranging from the personal factors to the properties of individuals related specifically to products and brand interactions. The concurrent validity reported for variables of interest in each of the studies considered have been classified by the criteria utilized. In turn, these reported have been assigned to one or more of the cells in Table 3. The reader will note that, in several cases, the concurrent validation activities of each author or groups of authors encompass multiple categories.

Illustrative Detailed Findings

Since much of the material which follows deals with research presented in summary form, it is useful to present three small recently completed segments of analysis from the 1970 Purdue Consumer Behavior Research Project (1970 PCBRP) to illustrate the general problems of criteria selection and measurement which appear throughout activity and attitude studies.

In the 1970 PCBRP, a number of questions were concerned with aspects of self-confidence, desire for change, risk taking, the personal appeal of a new product, and similar topics. These attitudinal questions were factor analyzed. One of the extracted factors was labelled "New-product Proneness." Four principle questions had factor loadings greater than .55.

I like to try things just because they are new.

I really enjoy beating everyone to a new product.

I often look for new things to try so I can stay ahead of my friends and acquaintances.

I investigate new products to make sure I am always using the most efficient ones.

Table 1 Summary of Characteristics of Studies

Project Title, and Investigators	Date & Place	Sample Size & Sex	Purpose	Dominate Types of Variables
Purdue Consumer Behavior Research Project: Philip Burger Edgar Pessemier, Dir. Richard Teach Douglas Tigert	June, 1966 Lafayette, Indiana	540* House- wives	Combined survey, laboratory, in-store and diary study of activity, attitude and purchase behavior for diverse classes of convenience goods.	Usage and brand purchasing Media exposure Attitude, interest & opinion (150) Personality Occupation capability Shopping Information/Risk Brand preference and attribute judgments Brand and slogan recognition
Purdue New Product Adoption Research Project: Indianapolis, Indiana Charles King, Dir. John Summers	Nov. 1967 Indianapolis Indiana	983F*	Broad study of general and product specific innovation, information and risk handling and adoption behavior for a wide range of consumer goods.	Daily activities Newness attitudes New product ownership Social activities & memberships Media exposure Psychological self-description Information sources Shopping Personal influence (self-designated opinion leadership)
Purdue Study of Diffusion of Architectural Innovations Charles King, Dir. Tom Ness	June 1968 Cook, Du Page & Lake Counties, Illinois	122 Architects	A study of the diffusion of architectural innovations among the architectural community, specifically identification and analysis of the innovator and in- fluential segments of the market.	Professional attitudes and values Personal attitudes and values Behavior patterns Personal self-description Self-designated innovativeness Self-designated opinion leader- ship
Purdue Study of Fashion Adoption Among College Students Charles King, Dir. Steve Baumgarten	Aug. 1968 Purdue Univ.	420F* 389M*	Broad study of fashion adoption among college students covering a wide range of products and styles.	Psychological Self-Descriptions (4 tests) Socio-political attitudes Media exposure Fashion adoption Local innovativeness Interpersonal communication

Chicago Life-style Study Douglas Tigert William Wells	Jan. 1969 U. S. Nat'l. (Market Facts)	829 F#	A substantial extension of the survey based life-style/psychographic component of the 1966 Purdue study using a national sample.	Attitude, interest and opinion (300) Product and brand usage Media Exposure
Canadian Life-style Study Douglas Tigert	June 1970 Canadian National (Burke Res.)	1849 F# 1042 M#	Adoption and extension of Chicago Life-style study to bilingual Canadian sample.	Attitude, interest and opinion (300) Value Media Exposure Product Usage & Ownership
1970 Purdue Consumer Behavior Research Project: Lafayette, Indiana F. S. DeBruicker Tom Hustad Edgar Pessemier, Dir.	Nov. 1970 Lafayette, Indiana	912 F* 912 M* (Husband and Wife Pairs)	Product class and brand specific extensions of 1966 Purdue Consumer Behavior Research Project to husband-wife pairs. Special attention to variables for market structure and market segmentation analysis.	Brand usage, preference and similarity judgments Brand attribute judgments - Product class contexts, benefits and AIOs Media exposure Product ownership profiles Information/ risk profiles Activity patterns Values Occupational capacity Consumerism attitudes Brand and slogan Rec.

*,# Data analyzed by Bruno was collected as part of a large commercial study. It included 60 AIO variables and sets of variables related to innovativeness, media exposure and persuasibility for 721 female subjects. Response time per subject to survey instruments ranged from slightly less than one hour to slightly more than four hours. Samples market with * payed to complete questionnaire. Gifts or the other rewards used as incentive for samples market #.

Table 2
A Taxonomy of Consumer Activity and Attitude Measures

Classes of Variables Characterizing the Individual Consumer		Classes of Marketing - Specific Direct and Mediating Variables			
	Marketing	Defining Personal Characteristics	Institutional	Product	Brand
Personal	Life Profile and Environmental Characteristics	Defining Socio-economic and personality characteristics	Availability and relative importance of various marketing institutions	Product ownership profiles	Brand ownership profiles
	Perceptions	Consumer's perceptions, knowledge, and beliefs about himself	Consumer's perceptions, knowledge, and beliefs with respect to marketing phenomena	Cognitive description at the product level	Cognitive brand level measures
Attitudes	Preferences	General value judgments and preferences	Consumer's sentiments about marketing as a process and about marketing institutions	Product class information at the preference level	Brand information at the preference level
	Behavior	Extent of consumer's actual participation in certain activities (generally mediating elements unaffected by marketing tactics)	Consumer's exposure to specific institutions and media	Product level behavioral dimensions related to product class usage	Brand level behavioral dimensions related to brand usage

Table 3

Classification of the Concurrent Validity of Selected Activity and Attitude Measures

Classes of Variables Characterizing the Individual Consumer		Classes of Marketing - Specific Direct and Mediating Variables			
Marketing		General	Institutional	Product	Brand
Personal	Life/Profile and Environmental Characteristics	Tigert (June 1966) Pessemier & Tigert (April 1970) King & Summers (Feb. 1971) Bruno (Aug. 1971)	Tigert (June 1966) King & Summers (Feb. 1971) Bruno (Aug. 1971)	Baumgarten (June 1971) Pessemier & Tigert (April 1970)	Pessemier, Burger, Teach, and Tigert (Feb. 1971)
		Summers (June 1968) Ness (August 1968) Baumgarten (June 1971)	Hustad & Pessemier (1972)		
Attitudes	Perceptions (Cognitive)	Summers (June 1968) Ness (August 1968) Baumgarten (June 1971)	Hustad & Pessemier (1972)		
	Preferences (Affective)	Baumgarten (June 1971) Pessemier & Tigert (April 1970)	Hustad & Pessemier (1972)		
Behavior		Summers (June 1968) Ness (Aug. 1968) Baumgarten (June 1971) Hustad (1972)	Tigert (June 1966) Bass, Pessemier & Tigert (Aug. 1969) King & Summers (Feb. 1971) Baumgarten (June 1971)	Hustad & Pessemier (1972)	Pessemier, Burger and Tigert (Nov. 1967) Pessemier, Burger Teach & Tigert (Feb. 1971)

A sum score variable called New-product Proneness was created from the responses to these questions.

In another part of the questionnaire a large number of recently introduced, innovative products were listed and respondents were asked to indicate ownership of the products. The ownership responses were summed by product types and for all products. These scores became the criterion for testing the concurrent validity of the New-product Proneness Index. Recall, the research started with a theoretical construct, new-product proneness. It was developed, measured and stated as a single index (for greater reliability and content validity). Next, it was tested by examining the degree to which it was correlated with relevant reported behavior. The results appear in Table 4 along with comparisons with other related variables of interest.

Table 4

Correlation of New-product Proneness Index With the New-product Use Index, Brand and Slogan Recognition Index and Selected Demographic Variables. (912 Female Subjects, 1970 Purdue Consumer Behavior Research Project)

Total <u>New-product Use Index</u>	.21***
New Food and Household N-p U I	.19***
New Clothing and Personal N-p U I	.18***
New Durable N-p U I	.11**
Total <u>Brand and Slogan Recognition Index</u>	.13***
Automobile B & SRI	.09**
Fast Food B & SRI	.10**
Tooth Paste B & SRI	.05
Number of Children at Home	.05
Wife's Education	.05
Family Income	.01
Significance Levels	
*** $\alpha \leq$.001	
** $\alpha \leq$.01	
* $\alpha \leq$.05	

It is reassuring to find that the New-product Proneness Index is a significant specific predictor of the associated behavior, new-product ownership. Of additional interest are the findings that basic demographic variables are not closely related to the New-product Proneness Index, but that the proneness index is associated with brand and slogan recognition. Although demographic and communication variables are significantly associated with some aspects of new product ownership, the proneness index makes a useful contribution to the challenge of identifying consumers who will tend to buy new and innovative products.

A second set of data illustrates a somewhat different situation. On an exploratory basis, self-designated occupation skill (SOS) variables were introduced in the 1966 PCBRP. Factor analysis of these variables yielded results which closely approximated the prior hypothesized relationships about common groups of self-designated occupation skills. No a priori hypotheses were conceptualized to identify potential relationships between either the raw variables or the factor (or sum) scores and various behavior criteria such as media exposure or product-class usage rates. On an exploratory basis, however, SOS variables were used as independent variables in various analyses. Subsequently, these variables were revised and additional hypotheses developed before a new SOS variable set was included in the 1969 Chicago Lifestyle Study. The data

from this study were factor analyzed, yielding results similar to those hypothesized. Also, the raw SOS variables were correlated with the readership of print media and the usage of various product classes. Again the findings were encouraging, motivating the use of a second revised set of occupation questions in the 1970 PCBRP survey. Factor analysis of these variables yielded satisfactory results and the raw variables have been used as predictors in several specific investigations (DeBruicker, in progress; Hustad, in progress). In addition, the occupational variables have been correlated with several large sets of potential criteria variables. Since it is impractical to present all of the relevant findings, here we will consider only the simple correlations between the raw self-designated occupation skill variables (as distinguished from occupational interest which was also studied) and four sets of criterion variables. The results of this test of concurrent validity are presented in Table 5.

Table 5

Correlations of 50 Self-designated Occupation Skill Variables With 66 Product Usage, 40 Product Ownership, 57 Magazine Readership and 77 Personal and Social Activities Variables. (912 Female Subjects, 1970 Purdue Consumer Behavior Research Project)

Criterion Variables	Total No. Correlations	Number of Correlation Coefficients	
		Sig. at $\alpha < .01$	Coefficients $> .1$
Product Usage Rates	3300	292	179
Product Ownership	2000	114	46
Magazine Readership	2850	217	116
Activities	3840	765	543
Total	11,990	1,388	884

Clearly, the number of significant correlations is more than 10 times the level one would expect by chance, but the SOS skill variables are more closely associated with participation in various activities than the directly market-related behavior such as usage, ownership, and readership. Of course, if the sets of usage and readership variables had been reduced such that the results were expected to be significant on the basis of the 1969 Chicago study findings, the performance of these measures could be further improved. In any event, these variables do describe a dimension of the consumer's character not easily tapped by other measures.

Although the relationships are relatively weak, SOS variables may help shape product, copy and media strategy. The reader is left to decide for himself whether evidence of the type cited above is adequate to judge the general validity of the variable set for scholarly and applied purposes. Recall, however, that even for scholarly efforts in our field of interest validity must ultimately be considered in terms of explicit market phenomenon. Clearly, the set of interesting criterion variables potentially related to SOS variables has not been exhausted. Finally, judgments about useful levels of validity cannot be made in the absence of a specific research question (and associated cost and efficiency issues).

A third type of validity issue can be illustrated by direct application of the SOS variables to a study of consumer attitudes toward marketing institutions and practices (Hustad & Pessemier, in press). Consumers were divided into pro, anti and uncommitted groups by their attitudes about business and marketing practices. Having identified the groups, the investigators were interested in the degree to which

Table 6

50 Self-designated Occupation Skill (SOS) Variables: Very Good (1) - Very Poor (6)
 Normalized Within Individual, Mean = 0, S. D. = 1.0
 (912 Female Subjects, 1970 Purdue Consumer Behavior Research Project)

Var. #	Variable Description	F*	Group Mean		
			Pro	Anti	Uncomm.
1	Doctor - General Practice	6.01	.48	.23	.51
2	Typist	3.84	-.83	-.58	-.71
4	Den Mother-Cub Scouts	6.95	-.66	-.37	-.70
9	Receptionist	4.11	-1.23	-1.05	-1.12
11	IBM Card Punch Operator	3.33	-.43	-.21	-.34
12	Salesgirl - Department Store	4.51	-.90	-.65	-.79
13	College Professor	9.27	.50	.17	.45
14	News Reporter - World Affairs	6.85	.53	.29	.54
19	Waitress - Restaurant	4.45	-.15	.03	-.26
20	Assembly Line Worker	5.71	.00	.19	-.16
23	Laboratory Technician	3.16	-.18	-.38	-.27
27	Psychologist	13.20	.23	-.17	.12
29	Beautician	6.08	-.29	.02	-.19
30	Bookkeeper	6.87	-.60	.27	.45
35	Professional Bowler	9.71	.36	.72	.53
39	Elementary School Teacher	4.69	-.30	-.39	-.50
42	Lawyer - General Practice	7.31	.54	.34	.59
43	Mathematics Teacher	3.34	.47	.68	.55
44	Guidance Counselor - High School	9.90	-.01	-.33	-.21
46	School Board Member	3.93	-.04	-.25	-.12
<u>11 Demographics</u>					
2	Occupational Status of Husband	4.82	3.31	2.41	2.02
9	Education	14.23	5.55	6.38	5.86

* F.05 = 3.00, F.01 = 4.61, F.001 = 6.91; 2, 909

members of the three groups differed in terms of their self-designated occupation skills (one of many classes of descriptor variables used to profile the three previously defined groups). The results for occupation variables (SOS) appear in Table 6.

Twenty out of fifty SOS variables were significant, six at the .001 level and six at the .01 level. Furthermore, when these variables were combined with a large set of additional independent variables in a discriminant analysis designed to predict group membership, the discriminant function correctly reclassified 53% of the holdout sample compared to about 33% expected by random assignment (Hustad & Pessemier, in press).

The principle point to note about this illustration is the problem of the criterion, group membership. The definition of pro, anti and uncommitted groups presents a validity problem in its own right. However, even if imperfectly defined, it is not unreasonable to expect personality traits associated with self-designated occupation skills to be associated with the differences in attitudes about business practices. Therefore, the concurrent test of validity is not devoid of theoretical support in terms of the two constructs, self-designated occupation skills and attitudes toward business.

The three examples presented in the preceding discussion illustrate some of the problems encountered in examining simple concurrent validity. The balance of this paper is devoted to reviewing the reported findings we have classified in Table 3 and have summarized in Table 7.

Synthesis of Findings

In the first study which we examine, Tigert utilized 37 factor scores derived from the responses of 344 individuals to 300 activity, interest, and opinion (AIO) measures. In this phase of his research, Tigert hypothesized the existence of dimensions that would be relatively independent of demographic measures. Thus, the validation criteria used by Tigert were applied in a negative sense--the lower the correlations, the more powerful the exclusive descriptive capacity of the activity and attitude measures.

Tigert found that many of the factor scores were, in fact, relatively independent of the demographic variables (Tigert, 1969).

The second study represented an attempt by Pessemier, Burger, and Tigert to discriminate among buyers of a new product utilizing numerous variables, including activity and attitude descriptors (Pessemier, Burger, & Tigert, 1967). In this study a buyer of new detergent was anyone who bought the product at least once in a seven-month period of diary keeping. An early buyer was one who purchased the product in the first 70 days after introduction. All remaining subjects who bought were classified as late buyers. The validation methodology employed was cross-classification analysis. Specifically, responses to a battery of measures were cross-classified against the early, late, and nonbuyers category.

The authors found that the significant variables were in four distinct categories: socioeconomic, trial-proneness, product-related, and informational. The following comments summarize the author's findings:

- 1) The opinion leadership question, the AIO factor scores on information seeking and risk avoidance, and the media factor scores on information seeking and risk avoidance, and the media factor scores were conspicuous for their inability to distinguish the kind of buyer.
- 2) Usage rate for the product class did not differentiate among early, late, and nonbuyer.
- 3) Nonbuyers had the least preference for the new brand, and the early buyer the greatest.
- 4) Early buyers were significantly less confident about their past brand purchases than late buyers, and late buyers were less confident than nonbuyers.

Table 7

Summary Descriptions of the Concurrent Validation Efforts Examined in this study

Investigator(s)	Classification of Attitude and/or Activity Measures Used	Purpose of Investigation	Results
Tigert (June 1966)	Life Profile and Environmental Characteristics general institutional	Contrast the profiles of users with non-users of various products and/or media and examine the market behavior of consumer types as defined by similar response patterns to AIO variables.	28 or 42 correlations between standard demographic measures and 37 AIO factor scores were significant at .05 level.
Pessemier, Burger and Tigert (Nov. 1967)	Behavior Institutional Behavior brand	Identify the characteristics of buyers and non-buyers of a branded laundry detergent to determine the differences between early, late, and non-buyers of the new brand.	Fourteen variables were significantly ($p < .05$) related to the kind of buyer in the cross-classification analysis.
Summers (June 1968)	Attitudes general Behavior general	Contrast profiles of fashion transmitters with non-transmitters along social activity, psychological characteristics, attitude and values, daily activity, and mass media exposure variables.	43 of 71 cross-classification comparisons were found to be significant at the .05 level.
Ness (August 1968)	Perceptions general Behavior general	Contrast profiles of self-designated architectural a. innovators vs. non-innovators b. information transmitters vs. non-transmitters c. attitudinal innovators vs. non-innovators d. perceived influentials vs. non-influentials	52 of 444 cross classification comparisons were significant at the .05 level.

Investigator(s)	Classification of attitude and/or Activity Measures Used	Purpose of Investigation	Results																						
<p>Bass, Pessemier, and Tigert (July 1969)</p>	<p>Behavior institutional</p>	<p>Expand the interpretation of media market segments through the use of market-related descriptors.</p>	<p>For four of the AIO factor scores, significant differences appeared between the mean values of at least two of the three clusters described by the authors.</p>																						
<p>Tigert (Aug. 1969)</p>	<p>Behavior institutional</p>	<p>Identify the psychographic characteristics of the audiences of selected magazines.</p>	<p>Top Ten Magazines Only</p> <p>Number of Correlation Significant at .01 level</p> <table border="1"> <thead> <tr> <th>Magazine</th> <th>Number of Correlation Significant at .01 level</th> </tr> </thead> <tbody> <tr><td>1. Playboy</td><td>136</td></tr> <tr><td>2. Business Week</td><td>121</td></tr> <tr><td>3. Time</td><td>115</td></tr> <tr><td>4. Nat'l Geographic</td><td>103</td></tr> <tr><td>5. New Yorker</td><td>102</td></tr> <tr><td>6. Life</td><td>89</td></tr> <tr><td>7. News Week</td><td>89</td></tr> <tr><td>8. Esquire</td><td>89</td></tr> <tr><td>9. Sports Illustrated</td><td>87</td></tr> <tr><td>10. Farm Journal</td><td>83</td></tr> </tbody> </table>	Magazine	Number of Correlation Significant at .01 level	1. Playboy	136	2. Business Week	121	3. Time	115	4. Nat'l Geographic	103	5. New Yorker	102	6. Life	89	7. News Week	89	8. Esquire	89	9. Sports Illustrated	87	10. Farm Journal	83
Magazine	Number of Correlation Significant at .01 level																								
1. Playboy	136																								
2. Business Week	121																								
3. Time	115																								
4. Nat'l Geographic	103																								
5. New Yorker	102																								
6. Life	89																								
7. News Week	89																								
8. Esquire	89																								
9. Sports Illustrated	87																								
10. Farm Journal	83																								
<p>King and Summers (February 1971)</p>	<p>Life Profile and Environmental Characteristic general institutional Behavior institutional</p>	<p>Expand the interpretation of media market segments through the use of market-related descriptors.</p>	<ol style="list-style-type: none"> 1. Television viewing was related to only 2 of 19 activity and attitude measures at .05 level. 2. Nine of nineteen activity and attitude measures were statistically significant for each of the magazine groupings. 3. Most of the activity and attitude measures reflected low intercorrelations with age, education, and income. 																						

Investigator(s)	Classification of Attitude and/or Activity Measures Used	Purpose of Investigation	Results
<p>Baumgarten (June 1971)</p>	<p>Life Profile and Environmental Characteristics</p> <p>Product</p> <p>Perceptions</p> <p>general</p> <p>Preferences</p> <p>general</p> <p>Behavior</p> <p>general</p> <p>institutional</p>	<p>Identify characteristics of opinion leaders among college students and degree of local innovativeness.</p>	<p>27 of 29 cross classification comparisons were significant at the .05 level.</p>
<p>Bruno (August 1971)</p>	<p>Life Profile and Environmental Characteristics</p> <p>general</p> <p>institutional</p>	<p>Identify the underlying dimensions of consumers' persuasibility and associate these dimensions with market-related behavior.</p>	<p>15 of 60 correlations were found to be significant at the .05 level.</p>

- 5) Early and late buyers indicated a greater willingness than nonbuyers to try new brands.
- 6) Early buyers see themselves as experimenters to a significantly greater degree than late or nonbuyers. However, early buyers did not perceive themselves as innovators.
- 7) Early buyers exhibit a higher degree of transmission of product information.
- 8) Late buyers, compared with early and nonbuyers, had a significantly higher educational level.

The third study was an extensive effort by Summers to identify women's clothing fashion transmitters (Summers, 1969). The focus was on the identification and profiling of the opinion leaders on interpersonal communication in the context of women's clothing fashions. The validation methodology employed was the cross-classification of various activity and attitude measures against fashion transmitter variables. Summers' conclusions are of interest:

"The results were consistent with previous research findings in the area of opinion leadership, and with few exceptions were also consistent with a priori judgments concerning women's clothing fashion transmitters. These findings demonstrate that real differences do exist between transmitters and nontransmitters, and that these differences can be measured.

"However, in the aggregate, these differences were not large. This indicates that, although concentrations of women's clothing fashion transmitters do exist and can be identified, opinion leadership in the area of women's clothing fashion is a wide-spread trait" (Summers, 1968).

A study conducted by Ness focused on isolating, identifying and profiling key change agents in the architectural diffusion process. Ness concluded that none of his measures very effectively differentiated between innovators and noninnovators (Ness, 1968).

In a taxonomic study based on magazine readership, Bass, Pessemier, and Tigert utilized attitude and activity measures to describe readership segments (Bass, Pessemier & Tigert, 1969). The four AIO factor scores significant at the .05 level were named by the authors as follows: compulsive, orderly housekeeper; careless, irresponsible; negative attitudes toward advertising's value; and outdoor, casual activist. In addition, the authors found that subjects in the three clusters exhibited significant differences in demographic, average usage rates on selected grocery products, and average scores on brand recognition.

Another study by Tigert was undertaken to show how advertising copy can be made to be more compatible with consumer life styles. (Tigert, 1969). In this analysis, 300 AIO characteristics were correlated with the reading levels of 53 individual magazines. Most of the correlations reported by Tigert fall in range of .10 - .19. The author attributes a partial explanation to the large number of responses which piled up in the nonreadership category.

In an article dealing with the relationship between activity and attitude measures and media exposure, King and Summers present empirical data relating personality and attitudes to the level of media exposure within six broad media classes and to exposure across the media classes. (King & Summers, 1971). Included in the six media classes was television viewing, as well as five magazine groupings: general interest mass appeal, news, home, women's general interest, and women's fashion. The measurements of personality and attitudes and values were based on a standardized personality inventory developed by the authors. Cross classification analysis was undertaken to determine the relationship existing between exposure to each of the six media groups and the activity and attitude measures.

- 1) The set of activity and attitude variables used were ineffective in identifying respondents highly exposed to television.
- 2) As a group, the activity and attitude measures used were effective in predicting magazine readership for all five magazine groups analyzed.
- 3) Some personality and attitude and values measures were more relevant to particular magazine groupings.
- 4) Most of the activity and attitude measures reflected low inter-correlations with age, education, and income.

More recently Baumgarten utilized activity and attitude measures to aid him in describing fashion adoption among male college students (Baumgarten, 1971). Baumgarten's conclusion are as follows:

- 1) In general the results indicate that opinion leaders are most significantly different from their followers along social activities, sociological, psychological and media exposure characteristics.
- 2) Demographics were found to be relatively poor indicators of both opinion leadership and innovativeness.

Bruno utilized activity and attitude measures to determine the relative usefulness of a persuasibility descriptor in a media evaluation model (Bruno, 1971).

In this research the author hypothesized the existence of characteristics which would be relatively independent of demographic measures and which would be linked to television exposure rates. Bruno found few strong relationships. Only one relationship, the correlation between self-esteem and age, was above .2. Only two correlations, the correlations between social activity and morning viewing and social activity and afternoon viewing were above .1.

A rudimentary, but illuminating summary of the studies cited in Table 7 appears below:

Significant Relationships Identified by Correlating
or Cross-classifying Activity and Attitude Variables
from Purdue Studies with Concurrent Validity Criterion
Variables

<u>Criterion Variable Classification</u>	<u>Results</u>
General	176 of 682*
Institutional	106 of 205
Product	47 of 49
Brand	34 of 40
Life Profile	101 of 189
Perceptions	142 of 564
Preferences	47 of 49
Behavior	246 of 737

*Read: 176 of 682 "mutually exclusive" correlations or cross classifications between activity and attitude measures and general concurrent validity criterion variables were found to be significant at the .05 level by the Purdue University researchers. The term "mutually exclusive" refers to the fact that a number of studies had multiple criterion variable classifications (see Table 3).

It is interesting to note that, for the Purdue studies considered in this paper, the proportion of significant product and brand specific direct and mediating variable relationships is higher than the proportion of significant general and institutional relationships identified. Moreover, of the individual consumer variables, preference bears the highest number of relationships to activity and attitude variables. The reader is cautioned to avoid seeking generalizable conclusions; however, because the studies tabulated have their origins in the same research tradition, because a number of the studies were

multiply classified, and because Ness' rather negative findings dominate the totals. Nevertheless, this type of summary, when considered with similar evidence from other research traditions, might provide useful insights into the applicability of certain types of activity and attitude measures.

Conclusions and Implications for Future Research

The preceding discussion provides an empirical backdrop against which we can cast some generalized conclusions and also identify areas requiring additional research.

First, it is clear that the activity and attitude research reviewed in this paper seldom incorporated explicit concurrent validity goals into the research design and subsequent analysis. Prior, general questions about validity appeared to be of secondary importance to the central focus of the author's(s') investigation. In many cases the expected relationships were not formally hypothesized. When explicit hypotheses were postulated, seldom was the extent of the expected relationship clearly described. This deficiency is probably due, in part, to the exploratory and empirical nature of the studies considered here. However, as empiricists we must be assured of the validity and reliability of measures which we utilized in describing consumer behavior. Hopefully, subsequent activity and attitude research will employ more extensively substantive a priori hypotheses and will document more clearly the extent of various predictive relationships indicated by empirical findings.

Second, the nature of the external criteria utilized for validation was often a source of concern. Occasionally, the criteria selected were arbitrary. Frequently the validity of the external criteria themselves was in doubt. On the whole, we conclude that the selection and analysis of validation criteria could be substantially improved.

Third, strong relationships were not identified in the studies reviewed here (most authors reported rather ambiguous results). This condition may be due to the fact that strong relationships between activity and attitude measures and certain external criteria do not exist. However, often it appeared that the criteria problem may have contributed to the dearth of strong identified relationships. Also, although many of the reported relationships were compatible with one or more theories of consumer behavior, there is little reason to believe that the links are strong and direct, especially for many interesting forms of low-salience classes of market behavior. We suggest that several justifiable forms of external criteria be considered in all future activity and attitude research and that the expected upper bounds of predictive power always be stated a priori. By considering alternative forms of the external criteria in this light, the validity and consequently, the potential usefulness of activity and attitude measures can be more conclusively verified.

We have stated elsewhere that standardizing some activity and attitude measures is a prerequisite to the development of efficient general purpose instruments. An equally important step should be a consolidation of the variety of approaches from diffuse exploratory studies. Efforts should be made to categorize the myriad of available measures as we have done here on a very modest scale. Our obvious goal is to build a set of activity and attitude measures which has wider application and acceptance due to its demonstrated linkage to various patterns of consumer behavior.

We are equally certain, however, that for many applied problems, activity and attitude measures that are specific to a product class are required. These variables take advantage of the specificity of the contexts, benefits, activities and attitudes that are typical of the brand strategy questions confronting marketing managers.

At this point, we will take an ambivalent position regarding the ultimate marketing usefulness of activity and attitude measures. After having considered the reliability of a selected set of them elsewhere (Bruno, 1971) and after examining their concurrent validity here, we are encouraged. However, much work remains to be done. We believe that it would be particularly helpful to draw generalizations from the results of large scale proprietary studies undertaken during the past five years. These empirical studies represent much larger and more comprehensive data bases than those conducted under the more modest auspices of the academic community.

Footnotes

1. University of Santa Clara
2. Purdue University

References

- Bass, Frank M., Edgar A. Pessemier, & Douglas J. Tigert. A taxonomy of Magazine Readership Applied to Problems in Marketing Strategy and Media Selection. Journal of Business, July 1969, 42.
- Baumgarten, Stephen. A Study of Fashion Adoption Among Male College Students. Ph.D. Thesis, Purdue University, June 1971.
- Bruno, Albert V. An Explicit Model for the Evaluation of Television Audiences. Ph.D. Thesis, Purdue University, August 1971.
- Bruno, Albert V. & Pessemier, Edgar A. An Empirical Investigation of Reliability and Stability of Selected Attitude and Activity Measures. Proceedings of 2nd Annual Conference of the Association for Consumer Research, University of Maryland, September 1971.
- DeBruicker, F. Stewart. Project Overview and Design Details. In Pessemier, DeBruicker and Hustad, The 1970 Purdue Consumer Behavior Research Project. Krannert Graduate School, Purdue University, June 1971.
- DeBruicker, F. Stewart. The Use of Product Class Specific, Situational, Personality, and Behavioral Measures in Market Segmentation. Ph.D. Thesis, Purdue University, (in progress).
- Guttman, Louis. Measurement as a Structural Theory. Psychometrika, December 1971, 329-347.
- Hustad, Thomas P. Information Handling Behavior for Consumer Types. Ph.D. Thesis, Purdue University, (in progress).
- Hustad, Thomas P. & Pessemier, Edgar A. Segmenting Consumers Market with Activity and Attitude Measures. Institute Paper No. 298, Krannert Graduate School of Industrial Management, Purdue University, March 1971.
- Hustad, Thomas P. & Pessemier, Edgar A. Will the Real Consumer-Activist Please Stand Up, Marketing Science Institute, in press.
- Kassarjian, Harold H. Personality and Consumer Behavior: A Review. Journal of Marketing Research, November, 1971, 409-419.
- Kerlinger, Fred N. Foundations of Behavioral Research. Holt, Rinehart and Winston. 1965.
- King, Charles W. & Summers, John O. Attitudes and Media Exposure. Journal of Business, 1971, 42.
- Ness, Thomas E. Change Agents in the Architectural Diffusion Process. Ph.D. Thesis, Purdue University, August 1968.
- Pessemier, Edgar A. & Tigert, Douglas J. Marketing Applications of Self-designated Occupation Skill Variables. Institute Paper No. 274, Krannert Graduate School of Industrial Administration, April 1970.

- Pessemier, Edgar A., Burger, Phillip, & Tigert, Douglas J. Can New Product Buyers Be Identified. Journal of Marketing Research, 1967, 4.
- Pessemier, Edgar A., Burger, Phillip, Teach, Richard & Tigert, Douglas. Using Laboratory Brand Preference Scales to Predict Consumer Brand Purchases. Management Science, February 1971, B371-B385.
- Summers, John O. The Identity of the Women's Clothing Fashion Transmitter. Ph.D. Thesis, Purdue University, June 1968.
- Tigert, Douglas J. A Psychographic Profile of Magazine Audiences: An Investigation of a Media's Climate. Working Paper, University of Chicago, August 1969.
- Tigert, Douglas J. Consumer Typologies and Market Behavior. Ph.D. Thesis, Purdue University, 1966.

AN OPERATIONAL CONSTRUCTION OF LIFE STYLE

Fred D. Reynolds and William R. Darden
University of Georgia¹

That the concept of life style and its operational referent psychographics are fashionable there is little doubt. The evidence is profuse! Major sessions are devoted to the topics at national and regional conferences; the number of papers and articles on the subjects has grown rapidly; two books have addressed the area exclusively; and there are tons of computer printouts stacked in the files of various marketing and advertising organizations which bear the psychographic imprint.

At first glance, the current status of life style and psychographics seems somewhat paradoxical. For the concept of life style is certainly not new, and yet it has been accorded a widespread acceptance in studies of consumer behavior only within the last decade. How can we account for this apparent reinvention of a concept and its escalating popularity?

It appears that several factors evolved in the marketing environment of the 1950's and 1960's and accumulated in a pattern which set the stage for the growing acceptance and use of life style and psychographics. These factors can be summarized as the motivation research legacy, the concept of an active audience,² increased computational modes, and an anticipatory notion of human behavior.

William Wells and Douglas Tigert (1971) recently presented the case for the legacy of motivation research. They pointed out that for all of its unworkability and controversial findings, motivation research brought people to the marketing research wasteland of percentages. Instead of describing an audience as

32.4 years old,
12.62 years of schooling,
90 percent married
with 2.1 children,

it became possible to think of

mothers who worried about getting the kids to school on time, old ladies whose feet hurt, . . .

and skinny kinds who secretly, but sincerely, believed that

The Breakfast of Champions had something to do with their batting averages.

The need to describe people in believable terms remained after motivation research lost its cynosure position. Indeed, the need became intensified as companies began to pour more and more oil to the flames of advertising and as consumer analysts became increasingly aware of the notion of an active audience.³ No longer could the consumer be construed as an inactive recipient in the communications process. No longer could we view the housewife as being pushed and pulled by the stimuli bombarding her from the mass media. Rather, the initiative of the audience had to be reckoned with, including the initiative of selectively attending to the increased promotional activity of consumer marketers. Moreover, as the notion of the active audience diffused, marketers became cognizant that people do not respond to the "real" situation but to the situation as they see it. Hence it became more important to "crawl into the skin of consumers" to see how they interpret and view situations and how their perceptions affect their behavior in the marketplace.

As the active audience capped the grave of motivation research, technological advances in computational modes were ushered into the research chapel via second and then third generation computers and more sophisticated software. With greater computational capabilities researchers became able to explore consumer behavior more fully using vast quantities of empirical data and multi-variate statistical techniques.

Then, there was the intuitively plausible notion going around that people didn't just buy products, but that they enter the market to replenish or to extend an assortment of goods "needed to support expected patterns of future behavior" (Alderson, 1965). Wrote Alderson, of course, was emphasizing the anticipatory nature of buying behavior and in so doing, he identified a key connection between purchase and the purpose of purchase. This connection was expressed by Alderson as:

Consuming habits are part of the pattern of living, but buying habits (purchasing habits and consumer habits) are only derived from this pattern. Buying habits . . . can be broken overnight with no real disruption in the pattern of living. Buying habits, in fact, can more safely be regarded as deliberately chosen routines designed to save time and energy for rational consideration of more important matters (Alderson, 1957:166).

In retrospect, it is not too surprising that consumer analysts, armed with greater data handling ability and the need to believably describe an active audience, used such a connection as a basis for quantitatively and qualitatively exploring extended bases for consumer description. Clark Wilson, for example, did exactly that. He said:

- (1) People . . . live according to established behavior and attitude patterns which can be identified and measured.
- (2) These Living Patterns, in turn, are related to other behaviors of more direct economic importance, such as product purchase and media exposure so that knowledge of such Living Patterns over and above knowledge of demographic characteristics, can be of economic value in marketing management (Wilson, 1966:306).

All of this, of course, is history. Now, life style and psychographics has arrived; it's with it! But the story is not all rosey. For one thing, in the apparent haste to put a measure on the market, researchers have done little conceptualizing about life style and psychographics. As Simmons (1971) observed, "My first and foremost impression about psychographics is that there is no general agreement as to just exactly what it is." Simmons offered his observation over five years after the first empirical papers appeared in the literature. Furthermore, he is correct (see Figure 1). And, as several commentators have recently noted, major problems often arise when research is conducted without the guidance of a conceptual orientation--researchers frequently produce poorly designed instruments, fail-to-find meaningful relations, and make faulty implications.

Figure 1

Alternative Definitions of Psychographics

Psychographics is a quantitative, multi-variate research procedure that gives numbers to common sense; like the fact that some people are more likely to be your customer than others; like the fact that some media are more likely to deliver the people you need than others. Psychographics could also be thought of as Measures of an Individual's Level of Expectation, a means for marketers to learn how to address the individual consumer as an individual through the medium of mass communications (Demby, 1971).

...a quantitative definition of the market based upon a systematically developed list of attitudes related to life-style and product benefits, constructed in such a way as to maximize product usage differential, against which advertising, marketing, and product decisions can be made (Heller as quoted in Hustad and Pessemier, 1971).

In its broadest sense, psychographics refers to any form of measurement or analysis of the consumer's mind which pinpoints how one thinks, feels, and reacts (Nelson, 1971).

...research that focuses on consumers' activities, interest, prejudices, and opinions. Variouslly called "psychographic" research, "life style" research and even (incorrectly). "attitude" research, it resembles motivation research in that a major aim is to draw recognizably human portraits of consumers. But it also resembles the tougher-minded more conventional research in that it is amenable to quantification and respectable samples (Wells and Tigert, 1971).

...the systematic use of relevant activity and attitude variables to quantitatively explore and explain the purchase and consumption of specific products or brands in the marketplace (Hustad and Pessemier, 1971).

...the term psychographics is derived from demographics and refers to psychological or 'Life Style' consumer characteristics (Tigert, 1969).

There is, then, a conceptual-empirical imbalance in the treatment of life style and psychographics. In another paper we addressed this imbalance in a theoretical and highly abstract manner (Reynolds and Darden, 1973). In this paper we seek to extend the address by presenting a more concrete conceptualization of psychographics; one that provides a meaningful organization to the phenomena of life style and one that allows for the possibility of deriving from it a set of stimuli that are relevant to the respondent in the research situation.

The paper is organized around three sections. The first demonstrates the potential disutility of conducting research without the guidance of an explicit conceptual framework; the second section presents the crucial component elements of psychographics; and in the final section each of the component elements are explicated and clarified in terms of its relation to psychographic research.

How to be Caught with your Pants Down Using Prior
Research in lieu of a Conceptual Framework

During the summer of '71 the first author examined the available literature on consumer out-of-town shopping behavior with the idea of conducting research and constructing a psychographic profile of the heavy outshopper. Because of the usual budget limitations, the questionnaire had to be restricted in length and would include only 85 general psychographic items. It was deemed imperative, then, to use meaningful and well tested items and scales. One way to approach this, of course, is to use the results of previous studies as a guide to build upon. And, to the extent that prior research is valid, this approach is useful.

One consistent finding was noted in each of the studies reviewed--consumers go out-of-town to buy clothes, particularly dress clothes, more than any other product (Hermann and Beik, 1968; and Thompson, 1971). It was assumed, somewhat naturally, that these up-scale housewives are relatively more fashion conscious than their peers and that they attempt to achieve and maintain a fashionable distinctiveness by shopping in larger metropolitan areas.

At the time, the findings of previous studies appeared to be a blessing without disguise. For, because of the fairly extensive fashion research we had completed, a set of ready-made, thoroughly tested psychographic scales were on hand each of which had been found related to the behavior of the female (and the male) clothes horse. These items were quickly integrated into the instrument and they accounted for about 30 percent of the total number of items included in the first section of the questionnaire.

The results? In an item by item contingency analysis, not one of the fashion related items significantly differentiated between frequent and infrequent outshoppers! Furthermore, in a three-way discriminant analysis using 28 summated scales as independent variables, the first fashion scale to appear in the rank ordering of the correlations between the discriminant function and the original variable was at position 17; needless to say, the scale was not significant in a one-way analysis of variance.

The reason?

The problem was not that of inaccurate scales. The usual factor analysis-subsample factor analysis was performed; the scales were structurally stable and highly consistent with structures uncovered in previous fashion research. Split-halves reliability coefficients were also congruent to previous studies. No, the scales were not too inaccurate; they were simply irrelevant to the phenomenon examined.

Part of the reasons for their irrelevancy is that the outshopper is not just a clothes horse. Rather, she tends to outshop for many types of shopping goods. She tends to be an active, cosmopolitan oriented person who enjoys shopping and is not time-conscious. Furthermore, she has a consistently poor image of her hometown shopping facilities.

Fortunately, the entire study was not lost; but that is another story.⁵ Nevertheless, the first section of the research instrument was, in part, wasteful and ineffective. Hopefully, the use of a viable conceptual framework for guidance would have precluded the use of irrelevant measures.

Life Style and Psychographics

This section presents a conceptual framework for life style and psychographics. The orientation is developed by first construing life style using George Kelly's theory of personal constructs as the theoretical base.

Life Style

What is life style? To get more precisely to the answer, consider the words separately. First, each person has a life and this makes sense only when viewed on a continuum of time; i.e., change. Persons are conceived, born, grow, mature, and decline over time, and we often refer to various aspects of this process in describing a person. But life is more than just change. If it were not so, there would be little, if any, difference between other parts of the universe and that part we call a person. As Kelly expressed it, life "involves an interesting relationship between parts of our universe wherein one part, the living creature, is able to bring himself around to represent another part, his environment." Life, then, has two aspects. It is measured on a time dimension and it has the capacity to retain its own identity while it represents other forms of reality. This interpretation of the word, life, comes from the philosophy of constructive alternativism, which is the point of departure for Kelly's theory. We can turn to the organization corollary of personal construct theory to give meaning to a person's style of life.

"Each person characteristically evolves, for his convenience in anticipating events, a construction system embracing ordinal relationships between constructs." Our emphasis here is on characteristically evolves a construction system. "Characteristically" notes the personalistic nature of the theory: "Not only are the constructs personal, but the hierarchical system into which they are arranged is personal too." Not only is the construction system personal, but it is also continually changing; it evolves. The system is viewed as relatively more stable than the individual constructs of which it is composed, but it nevertheless continues to take new forms. To realize that a person characteristically evolves a construction system is to recognize that he has a recognizable personality; one that, while relatively stable, is continually taking new shape.

By "construction system," Kelly noted that a system implies the grouping of elements in which incompatibilities and inconsistencies have been minimized. The elements are the personal constructs, the "transparent patterns" a person creates and lays over the events of his environment. They enable a person to chart a course of behavior; to live a "style" of life.

In essence, then, we view a person's life style to be the construction system that he characteristically evolves for himself. Since life style is considered to be the construction system, it is composed of construction subsystems each of which are made up entirely of personal constructs. From the dichotomy corollary of Kelly's theory we note that the construction system and hence its subsystems is "based upon constructs of constructs, concretistically pyramided or abstractly cross-referenced in a system of ordinal relationships." Yet, from the fragmentation corollary, we note that a person can employ his construction subsystems differentially and in an inferentially incompatible manner.

Thus, when persons allow us to notice their life styles, we never fully comprehend them. Rather, since a person's thinking is channelized and hence limited to a finite number of constructions at a time, we are usually allowed to glimpse, at best, only a few aspects of his life style. Noticing these repeatedly we tend to type him with some construct such as domineering or submissive, hard worker or lazy. Often, moreover, we notice inconsistent aspects of a person's style of life or a lack of pervasiveness in his personality, at least those aspects we have encountered. These points are important from the standpoint of the operational analysis of life style and we will return to them in the next section. First, however, we need to complete our

construction of life style by exploring, again constitutively, its expansion into a social aggregate situation.

For this purpose, we turn to the commonality corollary which states that to the extent that one person employs a construction of experience which is similar to that employed by another, his processes are psychologically similar to those of the other person.

This corollary provides the basis of construing life style in a cultural sense, but from the point of view of the individual person. In essence, we can develop a social aggregate whether it be dyad, triad, small group, or whatever by placing persons in the aggregate on the basis of the degree to which they construe their experience in the same way. This is to be sure often found when persons are grouped according to similarities in their upbringing and their environment. But the commonality of background does not guarantee that people will see things alike nor behave alike. Hence, in our construction of life styles we lay emphasis on the development of sub-cultures in terms of the members' similarities of construction systems.

Psychographics

Earlier we noted the definitional confusion of the word psychographics (see Figure 1). There is, however, a recurrent theme in the discussions of the term and that is the aim of understanding consumers, of drawing human portraits of them. This theme, in the language of the sociality corollary of personal construct theory reflects the desire on the part of researchers to "play a role" with consumers; i.e., being able to subsume the constructs of consumers with those we have created specifically for that purpose. In this light, we can view psychographics as the systematic operationalization of life style; i.e., constructs of persons' life styles. Thus, we seek to understand man as a person, as a process, and in order to do so, we attempt to subsume his construction system with our own. In attempting to analyze life style, however, we are not interested in subsuming it completely--even if that were feasible. Rather, we are interested in tapping the construction system for those constructs and subsystems relevant to consumer behavior, the product-related communicating, purchasing and consuming behaviors of persons. Figure 2 is a diagrammatic representation of this notion. The chain of nodes and links in the figure represent the construction system, albeit highly simplified, of a person. The broad line illustrates the subsystems within the total system that are relevant to this person's consumer behavior. The other aspects of the person's life style are irrelevant for our purposes. Indeed, any attempt to examine them in relation to the consumer relevant aspects of the system would tend to produce inferentially incompatible results. It is only when we view consumer, non-consumer subsystems in light of the superordinate systems within which they operate that they cease to become incompatible. This, we believe, to be one of the main problems with the use of standardized clinical measures in many previous attempts to predict consumer behavior. It is not that the measures are inaccurate per se; it is simply that such measures are subsuming irrelevant aspects of the construction system and hence are incompatible with consumer behavior.

Thus far, we have implied that psychographics is the systematic operationalization of life style. And we have posited our focus of convenience as the consumer realms of behavior. In other words, we do not narrowly construe psychographics as we do not narrowly construe life style. But, we must be careful to delineate the limited ranges of convenience of our systems in order to (hopefully) lessen the possibility of obtaining incompatible results. This end, of course, is ideal and perhaps not concretely attainable.

Figure 2

Abstract Representation of a Person's
Construction System



Source: Reynolds and Darden (1973).

Nevertheless, the continual striving towards the goal is necessary for the viability of a research tradition. This raises the question of what types of constructs should we consider regnant for the operation of life style within the focus of convenience we have posited.

The major types are posited in the following definition which is an elaboration of the one by Hustad and Pessemier (see Figure 1):

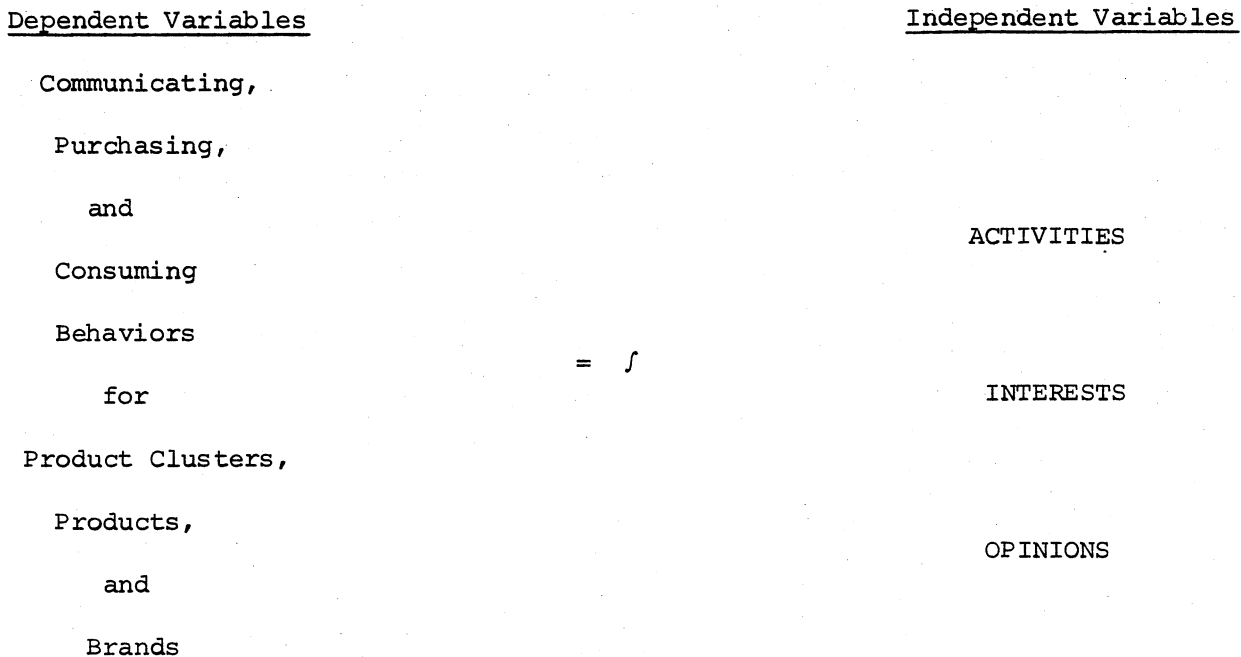
Psychographics is the systematic use of relevant activity, interest, and opinion constructs to quantitatively explore and explain the communicating, purchasing, and consuming behaviors of persons for brands, products and clusters of products.

Diagrammatically, the component elements of the definition are shown in Figure 3. It should be noted that the dependent variables are also a part of persons' life styles. Each, then, might for some purposes be used as independent variables. In the main, however, these are viewed as instrumental behaviors engaged in to fulfill other aspects of persons' life styles. Moreover, they represent behaviors which, when conducted in certain ways but not others, are determinant to the success of a firm's market offering.

The independent variables are, of course, the psychographic constructs. The question, to which we now turn, is what set of AIO constructs is most likely to guide research toward fruitful ends?

Figure 3

Dependent and Independent Variables

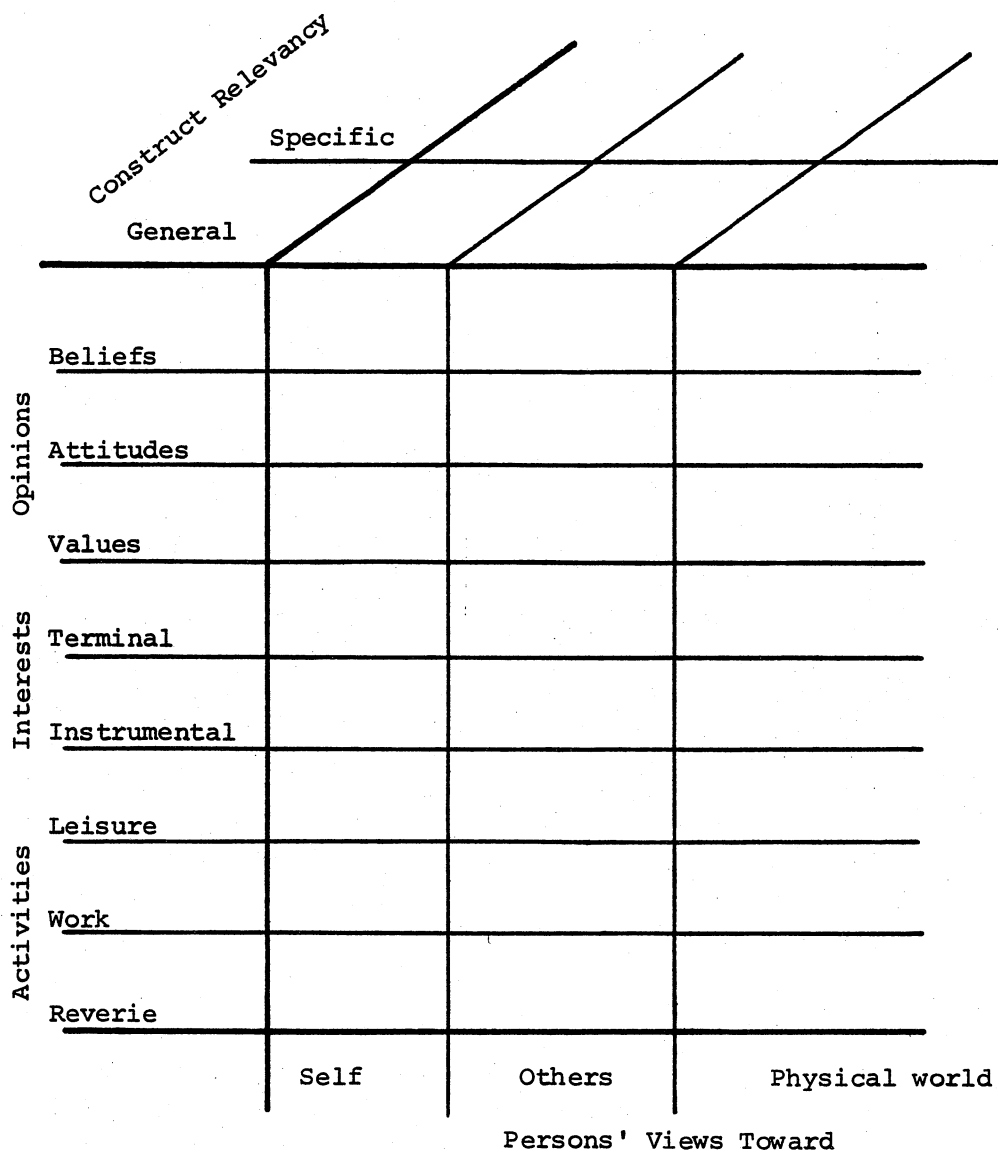


AIO's: An Elaboration

Figure 4 depicts the major psychographic categories along with two other dimensions of consideration in the generation of a research instrument.

Figure 4

Major Psychographic Categories

AIO's

In the figure, only the major categories of AIO constructs are depicted. Each of these, however, can be specified further in an expanding tree diagram when necessary. For guidance purposes, one can use the major categories and then elaborate on them in terms of the problem at hand.

An activity is a manifest action and as such is usually observable (Hustad and Pessemier, 1971). Moreover, it can be viewed as an episode; i.e., as having a beginning and an end. The activities of people, then, are a series

of episodes, connected end-to-end during a time period (Foote, 1966). The major categories of activities are leisure, work and reverie. Although each of these may be important for predicting certain types of behaviors, it appears that most psychographic research has focused upon leisure time and housework (Wind, 1971, and Wells, 1971).

An interest in some object, event, or topic is the degree of excitement that accompanies both special and continuing attention to it (Reynolds and Darden, 1972-73). Interests can be viewed in terms of two major categories, instrumental and terminal. An instrumental interest is one in which the topic of interest is viewed as a means to an end whereas a terminal interest is viewed as an end rather than a means. Much of the published works on interests have been of the instrumental kind, particularly vocational interests. More recent psychographic studies have explored, however, what a priori appear to be terminal interests such as fashion, child care, and homemaking. There is a continuing need, however, for research to determine the natural classifications of interests. Community mindedness, for example, might be a terminal interest construct for some persons and an instrumental one for others.

An opinion is a verbal or written 'answer' that a person gives in response to stimulus situations in which some 'question' is raised. It is used to describe interpretations, expectations, and evaluations--such as beliefs about the intentions of other people, anticipations concerning future events, and appraisals of the rewarding or punishing consequences of alternative courses of action (Hovland, et al., 1953).

In one sense, most psychographic research can be typed as opinion research because of the nature of data collection--self report rather than direct observation. In a more appropriate sense, and one that is also congruent with the above definition, an opinion can be viewed as the expression of a belief, attitude, or value (Rokeach, 1968). Beliefs, attitudes and values are of course, hypothetical constructs and hence unobservable. They must be inferred from stated opinions or other behaviors. Nevertheless, the belief and value systems surrounding objects and situations are important aspects of psychographic research. Furthermore, there is an extensive body of theoretical and empirical literature in this area; particularly noteworthy is the work of Rokeach and his colleagues (1968^a, 1968^b, 1968-69) which appears to be highly consistent to the theory of personal constructs.

Persons' Views

In the discussion of life style it was noted that people have the capacity to retain their own identities while representing other forms of reality. This suggests that researchers consider persons' views toward themselves, toward other persons, and toward aspects of situations (the physical world). Much of the AIO work published, however, emphasizes views toward self (Wind, 1971). There is, then, the need to continue exploring other aspects of persons' views.

Construct Relevancy

"Relevance" implies a traceable connection and a significant one in the sense that it contributes to the understanding of the matter at hand. Thus, in commenting on the relevancy of psychographic constructs, we are interested in the degree to which such constructs contribute to our understanding and prediction of consumer behavior. In examining their relevancy we are not addressing reliability, validity, etc. in a measurement sense; rather the

the central issue deals with the generalized--specificity continuum that has been noted by Ziff (1971) and Hustad and Pessemier (1971).

Basically, the general-specific continuum refers to the degree to which the operations employed in a psychographic study approach product or brand specific attributes. Table 1 illustrates the nature of measures approaching each end of the continuum as well as an intermediary position. The nature of these measures is clear. The first opinion leadership scale makes no mention of products or product classes, and is, therefore, general. The second scale moves to a more specific nature in that it is composed of statements related to the topic area of clothing fashions. The third scale illustrates further movement toward specificity in that consumer communicating activities are assessed toward a specific product, the midi.

Some important consequences follow the option taken as to the nature of the measures employed. Primary among these are the costs and flexibility of a given psychographic instrument. In general, the greater the generality of the measures the less costly and more flexible the instrument. Generalized measures are relatively less costly in the sense that they can be used as a standardized battery of measures repeatedly across a variety of contexts without redevelopment. Greater flexibility is achieved with general measures since a single psychographic instrument can be used to explore AIO relationships for a vast number of products and brands. Moving towards the specific end of the continuum, however, increases the developmental costs and reduces the flexibility of the instrument. As Ziff (1971) expressed it: "using product-oriented or benefit variables infers the necessity of tailoring the psychographic instrument to an individual product, or at least to a class of products similar in nature."

From the standpoint of costs and flexibility it would appear that generalized interest and activity measures hold an inherent advantage over more specific measures. Yet, our criterion of relevancy is not necessarily equated with the costs and flexibility of an instrument. Rather it is the degree to which our understanding and prediction of consumer behavior is enhanced.

Two studies have sought to examine the relevancy question empirically. In the first, Ziff (1971) concluded that the more product specific measures yielded relatively more insight than did general measures: ". . . the individual product segmentations, by virtue of their concentration on attitudes relevant to a particular class of products, provided a much deeper understanding of the values and needs in a product class than did the overall segmentation." Yet, given the greater insight obtained from specific measures, Ziff also found that the more general measures yielded identifiable and understandable psychographic segments of the drug market.

In the second study, Reynolds and Darden (1972_a) found that specific measures were better predictors, statistically, than general measures. Nevertheless, they found that general measures yielded insight into the nature of the behavior not found in the specific measures. They concluded that general measures should continue to be combined with specific measures in psychographic research. This approach allows researchers to maintain sufficient flexibility and prediction while lessening costs of instrument development.

In addition to the empirical quest for the answer to the relevancy question, the theory of personal constructs offers some useful guidelines.

The range corollary states that a personal construct is convenient for the anticipation of a finite range of events only. This implies that consumers have few, if any, constructs which are universally relevant for all types of products or other events. This corollary, of course, emphasizes the need for specific measures which, as noted, is also supported by existing research. Persons may construe products, however, in a sense not determined by traditional product classifications. We might find, for example, more similar

Table 1

Example Measures Illustrating the General-Specific Continuum of Psychographic Constructs

-----GENERAL-----		-----SPECIFIC-----	
Self-Designated Opinion Leader	Self-Designated Fashion Opinion Leader	Self-Designated Midi Opinion Leader	
1. My friends or neighbors often come to me for advice.	1. My friends and neighbors often ask my advice about clothing.	1. Several of my friends asked my advice about whether the midi would become a fashion or not.	
2. I sometimes influence what my friends buy.	2. I sometimes influence the types of clothes my friends buy.	2. I told my friends that the midi was a recurrent style from yesteryear and that it would be unflattering and make most women look older.	
3. People come to me more often than I go to them for information about brand.	3. My friends come to me more often than I go to them for information about clothes.	3. At coffee breaks my friends tended to ask my opinion of the midi more often than I initiated the conversation about it.	
	4. I feel that I am generally regarded by my friends and neighbors as a good source of advice about clothing fashions.		
	5. I can think of at least two people whom I have told about some clothing fashion in the last six months.		

Source: Reynolds and Darden (1973).

constructions for clothing items and electrical appliances of high social visibility. Thus, while persons use certain constructs for a finite range of products, these constructs often cover a range, and this range is not necessarily (although in some instances it may be) synonymous with traditional product classifications such as electrical appliances, furniture, cosmetics, drugs, etc. The import seems clear: use product specific or benefit type measures for products construed by consumers only in terms of benefits and use more general measures that allow us to subsume construct systems within whose ranges clusters of products fall.

In addition, there is the managerial situation involving the introduction of new products and/or suggesting new constructs to consumers with which to differentiate existing products. In these situations the modulation corollary of personal construct theory sheds insight since, in the assumptive structure of the theory, it establishes the conditions which governs the change within a person's construction system. In essence, the modulation corollary tells us that if we wish to change some aspect of a person's life style we must operate within his permeable superordinate constructions or the broader aspects of his life style. And, although it has not been demonstrated, we suspect that the more general psychographic measures come closer to subsuming consumers' superordinate subsystems than specific measures. Rokeach's (1968^b) discussion of value systems implies this also. If this is so, the use of a relatively large number of general measures may prove more helpful in situations where we desire to change some aspect of consumer behavior. This needs exploration, however, in terms of the superordinate-subordinate relationships of psychographic measures and their epistemology with consumer construction systems.

Conclusion

This paper represents an attempt to develop an operational construction of life style, one that provides a meaningful organization to the phenomena and one that allows the possibility of deriving from the framework a set of stimuli that are relevant to respondents in research situations.

Since the framework specifies only the crucial component elements of life style analysis, there is much work yet to be done. Hopefully, this paper will stimulate additional formulations and research addressed to the questions raised in the paper.

Footnotes

1. Assistant Professor of Marketing and Associate Professor of Marketing, College of Business Administration, University of Georgia.
2. These comments are adapted from Reynolds (1972).
3. For a review of the notion of the active audience, see Cox (1967).
4. See Reynolds and Darden (1971, 1972^b) and Darden and Reynolds (1971 and 1972).
5. Reynolds and Darden (1972^a).

6. Kelly's theory is concerned with the way persons conceptually organize and structure their environments and the processes by which they change their conceptual structures. The theory is formally stated as a Fundamental Postulate and eleven Elaborative Corollaries. This paper discusses only certain corollaries; for full discussions, see Kelly (1955), Bannister (1962), and Bannister and Mair (1968).
7. For a more complete discussion of the dependent variables used in the analysis of consumer behavior, see Reynolds (1972).

References

- Alderson, Wroe. Marketing Behavior and Executive Action. Homewood, Ill.: Richard D. Irwin, Inc., 1957.
- Alderson, Wroe. Dynamic Marketing Behavior. Homewood, Ill.: Richard D. Irwin, Inc., 1965.
- Bannister, D. Personal Construct Theory: A Summary and Experimental Paradigm. Acta Psychologica, 1962, 20, 104-20.
- Bannister D. & J. M. M. Mair. The Evaluation of Personal Constructs, London: Academic Press, 1968.
- Cox, Donald F. (ed.), Risk Taking and Information Handling In Consumer Behavior, Boston: Graduate School of Business, Harvard University, 1967.
- Darden, William R. & Fred D. Reynolds. Shopping Orientations and Product Usage Rates, Journal of Marketing Research, 1971, 8, 505-8.
- Darden, William R. & Fred D. Reynolds. Predicting Opinion Leadership for Men's Apparel Fashions. Journal of Marketing Research, 1972, 9, 324-8.
- Demby, Emanuel H. Psychographics: Who, What, Why, When, Where, and How. In King and Tigert, Attitude Research Reaches New Heights. Chicago: American Marketing Association, 1971.
- Foote, Nelson F. The Time Dimension and Consumer Behavior. In J. W. Newman (ed.), On Knowing the Consumer. New York: John Wiley and Sons, Inc., 1966.
- Hermann, Robert O. & Leland L. Beik. Shoppers' Movements Outside Their Local Retail Area, Journal of Marketing, 1968, 32, 45-51.
- Hovland, Carl I., Irving L. Janis, & Harold H. Kelley. Communication and Persuasion. New Haven and London, Conn.: Yale University Press, 1953.
- Hustad, Thomas P. & Edgar A. Pessemier. Segmenting Consumer Markets with Activity and Attitude Measures. Unpublished working paper, The Herman C. Krannert Graduate School of Industrial Administration, Purdue University 1971.
- Kelly, George A. The Psychology of Personal Constructs. Vol. I, II, New York: N. W. Norton & Co., 1955.
- King, Charles W. & Douglas J. Tigert, (eds.), Attitude Research Reaches New Heights. Chicago: American Marketing Association, 1971.
- Nelson, Alan R. Psyching Psychographics: A Look at Why People Buy. In King and Tigert, Attitude Research Reaches New Heights. Chicago: American Marketing Association, 1971.
- Reynolds, Fred D. Psychographics: New Dimension in Consumer Analysis. Athens, Ga.: Division of Research, College of Business Administration, University of Georgia, 1972.
- Reynolds, Fred D. & William R. Darden. Mutually Adaptive Effects of Interpersonal Communication. Journal of Marketing Research, 1971, 8, 449-54.
- Reynolds, Fred D. & William R. Darden. Intermarket Patronage: A Psychographic Study of Consumer Outshoppers. Journal of Marketing, 1972a, 36.
- Reynolds, Fred D. & William R. Darden. Predicting Opinion Leadership for Women's Clothing Fashions. Combined Proceedings, American Marketing Association, 1972, b.

- Reynolds, Fred D. and William R. Darden. An Analysis of Selected Factors Associated with the Adoption of New Products. Mississippi Valley Journal of Business and Economics, (Winter, 1972-73), forthcoming.
- Reynolds, Fred D. & William R. Darden. Construing Life Style and Psychographics. In William D. Wells (ed.), Life Style and Psychographics. Chicago: American Marketing Association, forthcoming, 1973.
- Rokeach, Milton. Beliefs, Attitudes, and Values. San Francisco: Jossey-Bass, Inc., 1968a.
- Rokeach, Milton. A Theory of Organization and Change Within Value-Attitude Systems. Journal of Social Issues, 1968b, 24, 13-33.
- Rokeach, Milton. The Role of Values in Public Opinion Research. Public Opinion Quarterly, (Winter 1968-69) 32, 547-59.
- Simmons, W. R. Overall Impressions of Psychographics. In King and Tigert, Attitude Research Reaches New Heights. Chicago: American Marketing Association, 1971.
- Thompson, John R. Characteristics and Behavior of Outshopping Consumers. Journal of Retailing, 1971, 47, 70-80.
- Tigert, Douglas J. Psychographics: A Test Retest Reliability Analysis. Proceedings. Fall Conference of the American Marketing Association, 1969, 310-15.
- Wells, William D. AIO Item Inventory. Unpublished paper, Graduate School of Business, University of Chicago, 1971.
- Wells, William D. & Douglas J. Tigert. Activities, Interest, and Opinions. Journal of Advertising Research. (August 1971), 27-35.
- Wilson, Clark L. Homemaker Living Patterns and Marketplace Behavior--a Psychometric Approach. Proceedings. World Congress of the American Marketing Association, 1966, 305-31.
- Wind, Jerry. Life Style Analysis: A New Approach. Combined Proceedings. American Marketing Association, 1971, 302-5.
- Ziff, Ruth. Psychographics for Market Segmentation. Journal of Advertising Research, (April 1971), 3-10.

ANALYSIS OF LIFE STYLES FOR STUDY
IN CONSUMER BEHAVIOR

Flora L. Williams¹
Purdue University

Statement of the Problem

The term life style is used casually to point out differences in the way people live. Life style is assumed to be an important variable for studying consumer behavior. Desired life style determines consumer choice and the choice is a reflection of life style.

The term life styles conveys different meanings in common usage, but should not be ambiguous for research purposes. The purposes of this study were to define life style for operational use in research, to select indicators of life style, to analyze factors that contribute to life style so that changes in consumer behavior can be understood, to describe the factors as to their contributions and limitations for researching life styles.

Before going to empirical data for rigorous examinations of the correspondence between types, one should decide the relevant data and a limited number of complete cases, and whether all the attributes included in the type are needed. Components in life style are conceptually distinct entities. Operationally maintaining this distinction is important in research.

Methodology

The nature of science requires that, first, observations and then categorizations, comparisons and classifications are made. Research organizes as well as acquires knowledge.

The method for conceptualizing life styles for analysis purposes was primarily a survey of the literature and research pertinent to the subject. Factors found were studied for commonalities from various sources. Typologies and a theoretical model were developed. Six case studies were selected and revised from about 100 case studies that seemed to be representative of life style types. A panel of judges selected the six for purposes of analysis. These case studies were analyzed using the indicators selected for the operational definitions of life style.

Hypothesis

The statement of hypothesis theorized what the components of life style were. Components of life style were described as commodities and services used, the manner in which they are used, satisfactions obtained, pattern of use of all resources available, and attitude about the level of living obtained. The pattern of resource use includes decision making, production in and out of the home, leisure time activity, and types of consumer behavior. These components are those sometimes considered in defining level of living. The hypothesis was used as a benchwork definition for selecting indicators for analyzing life styles. Components as stated were tested for their ability to be used as indicators.

Review of Literature and Research

The North Central Regional Research Project, NC90, entitled "Factors Affecting Patterns of Living in Disadvantaged Families" begun in 1966 and funded by the Cooperative State Research Service of the U. S. Department of Agriculture, had for one of its purposes: select and/or construct and evaluate indicators, indices, and measurements of patterns of living. The data is currently being analyzed.

The factors selected in the "pattern of living" concept for determining contributions to differences in families at a given income level and family size were:

- (1) mobility, geographic and intergenerational (education and occupation);
- (2) attitude toward situation being better or worse off, financial, living conditions, job opportunities, and opportunity for children;
- (3) value orientations, abstractness-concreteness, control-fatalism, equalitarianism-authoritarianism, integration-alienation;
- (4) housing, own or rent, number of rooms, hot and cold piped water, toilet and bath, telephone, garbage collection, television, and newspaper;
- (5) transportation, means, distances traveled to work, and problems;
- (6) relatives and family interaction; and
- (7) resource procurement and use, sources of income, size of income, financial problems, fixed commitments, and perceived adequacy of income.

In an article by Emanuel H. Demby entitled "Over-the-Counter Life Style" (1972), the predictive powers of demographic data are questioned. The researcher found very different attitudes about the amount of money persons earn which affected purchasing power. The factor he found discriminated persons was their ability to visualize experiences beyond the immediate environment and integrate different products into their life style. A method of life style research had been developed called psychographics which "seeks to measure (1) how appropriate to their life styles people's purchases are; (2) how products, political and social activities relate to the life style and individual self-concept; and (3) what is lacking in a person's life-style or self-concept that would produce a desire to buy a product or participate in a given activity."

The importance of attitudes in influencing behavior has been indicated by a variety of studies. A consumer behavior study found that "attempts at persuasion contradicting an individual's predisposition toward a given topic usually will be selectively perceived (Engel, 1972). Selective perception takes the form of selective exposure (ignoring the stimulus), selective distortion (missing the point of the message), and selective retention (no conscious recall of the message)." The study reported the following finding:

A company attempted to increase sales of raw berries. Their survey found market segments of (1) heavy users who valued tradition, (2) light to moderate users who were ambivalent about the role of homemaker, and (3) non-users described as "emancipated women" who avoided products connoting

tradition. Advertising tried to persuade the ~~non~~-users to buy the berries for the traditional sauce but with no success since the product contradicted their basic attitudes. Success was achieved by selling the berry juice for use in mixed drinks.

Certain influences will or will not affect life styles, therefore, depending upon the consumer's attitude, value orientation, or philosophy which are usually changed very slowly.

Influences upon life style that can be identified for comprehending how consumer behavior is changed are social ideals, educational level, occupation, income, goods available, individual differences, residence and aspirations. Environmental influences, furthermore, include social arrangements such as custom, advertising, governmental regulation and provisions, new events in society, and associations.

Practices of the social group to which individuals and families belong (since basic needs are met in that group) or are considered just above it, become the styles considered by the young child. These influences of social arrangements are somewhat subtle and may not be readily realized by the individual concerned.

Religious or cultural ideals mold styles. Practices of the group to which families belong may dominate the life style so much that a subculture persists. A cult of simplicity may resist the material aspects of living of the society at large. Families may form a commune to achieve their social ideals through cooperation in production and consumption. In striving to find patterns that better fulfill their ideals, alternate styles to the main theme of the American way of life are experienced. Strains of ethnic group behavior are observed in individuals adding to variation in life styles.

Size of income and price of goods at a given time influence life styles as illustrated by the City Worker's Budgets compiled by the Department of Labor. All three budgets provide for "maintenance of health and social well-being, the nurture of children, and participation in community activities." They represent the necessary conventional and social as well as biological requirements of a self-supporting family. But the lower budget family is limited to rental housing without air conditioning. It relies on public transportation and depends on free community facilities for much of its recreation. At the other extreme, the life style for the higher budget families means home ownership for most families, a relatively new car, significantly more and better household appliances and equipment, and a greater amount of paid-for services and entertainment.

Education is a dominant force as it is a means by which life styles are passed on from person to person or newly constructed for the individual. Education not only increases awareness of alternative styles of life, but provides means for changing one's style. A change in what is satisfactory frequently occurs with increased education. An emphasis may be placed more on aesthetic expressions than material goods.

Individual differences account for variations in styles of similar income groups or groups with similar value systems. Age, stage of family life cycle, sex, physical and mental characteristics, abilities, vitality, interests, and the specific knowledge which experiences and the interpretation of

their meaning have brought, develop unique styles. Early experiences may teach a child whether or not he can manipulate his environment by how quickly his cries are attended to. Rewards for successful efforts encourage further attempts.

There are other factors such as war, unemployment, geographical conditions and natural forces, political arrangements, government provisions and technological developments that directly influence life styles.

Development of Typologies

Several typologies were collected and developed for analysis of life styles. Typology is the study of types. In research, the type is merely a tool for the ordering of concrete phenomena. It is an imagined entity, created only because through it reality may be more easily understood. The typological structure organizes and sets in words descriptions of those characteristics which certain phenomena have in common. The general form of those phenomena become visible by means of the typological methods which may not have become apparent by other means of measurement. Study and implications may then ensue. The type may be composed of essential characteristics or one fundamental characteristic. Therefore, a type may be a system or a general denominator for a set of phenomena. The type is not just a simple unit but is a multiple of units.

In summary, the type is a "purposive, planned selection, abstraction, combination and accentuation of a set of criteria with empirical referents that serve as a basis for comparison of empirical cases" (McKinney, 1966).

The types of physical and psychological needs have been described by many writers. They are recognized as dominating forces of life style since they are motivation for human activities designed to fulfill goods. A listing of types does not explain the nature of needs. Various needs are not equal in their driving force. One type dominates life style depending on the state of satisfaction of other desires. Maslow suggests that:

"the chief principle of organization in human motivational life is the arrangement of needs in a hierarchy of less or greater priority or potency. The chief dynamic principle animating this organization is the emergence of less potent needs upon gratification of the more potent ones. Physiological needs, when unsatisfied, dominate the organism, pressing all capacities into their service and organizing these capacities so that they may be most efficient in this service. Relative gratification submerges them and allows the next higher set of needs in the hierarchy to emerge, dominate, and organize the personality. Instead of being hunger obsessed, it now becomes safety obsessed. The principle is the same for other sets of needs in the hierarchy, love, esteem, and self-actualization (Maslow, 1954).

Life styles are observed in the arrangements of house design and furnishings. Families may be labeled as: (1) economy-centered, (2) family-centered, (3) freedom or aesthetic-centered, and (4) prestige-centered on the basis of their housing decisions (Beyer, 1965).

A set of values guides behavior and tends to make it uniform so that life styles can be described as typologies. A typology was constructed to study values underlying family decisions of a middle-class sample (Schlater, 1969). The types used for analyzing responses were traditional, social, autonomous and change-prone.

To the traditional homemaker, the husband's position is considered most important, wife's position next and children's last. Children should be taught to behave. Husband's job is very important to the family. Materials such as car, furniture and money are of great importance to the family. Wife should have high standards of task performance. Categorical statements indicate the following of prescriptive behavior. There may be little control of the adult child. Security is more important than achievement or risk. Roles are probably clearly defined. The traditional type emphasizes production, duty, rights and responsibilities, and other-direction.

The homemaker interested in the social interaction perceives the wife's first responsibility as the loving care of her family. She is not free to decide what she will do. She does what her family needs and wants her to do. Husband's job is important to provide things for the family. He should like his job. Materials are important as they contribute to the family as a whole. New furniture and car are not mandatory. There may be categorical statement and conditional ends such as family needs and desires that dictate decisions. Many decisions are made by the family together. Responses of people are emphasized. The social type emphasizes affection, love, solidarity, loyalty, and other-direction.

For the autonomous-rational homemaker, time is free to be used in ways she sees beneficial to herself and family. The emphasis is on individual independence and growth as well as relative equality in decision making. There are more conditional and logical statements than categorical statements. Consequences of possible choices may be discussed. New car or furniture are seen in terms of its contributions to members of the family. Children's behavior is understood and reasonable control advocated. Wife has freedom to manage as she perceives best. The autonomous type emphasizes growth and development, fairness, impartiality, and responsible inner-direction.

The change-prone homemaker is described as doing what she feels like and would enjoy. In this family the new car and new furniture are desired for the fun of having it. The family would move if the husband wants to because it would be good to have a new experience. There are few categorical statements. Sometimes specific action statements are made. Children's behavior is accepted as self-expressive. Individuality is highly prized. There may be no attempt to control older children. Few prescriptions are handed out. The roles family members hold are probably indefinite and changing. The change-prone type emphasizes the new, the novel, and inner-direction. Little thought is given to consequences.

In more recent research investigating decision-making and value orientations of Guatemalan peasant families, the traditional orientation was described as stability and the type fatalism was added. Families in this other culture could be described by the five types (Baker, 1972).

When an individual's philosophy is known, his life style or consumer behavior may be understood. This includes attitudes toward and habitual evaluations of goods. When individuals have conflicting philosophical perspectives, agreement on resource use is difficult. Some philosophical perspectives which guide activities are: the new is better than the old, the old is better than the new, self-interests and concerns are most important; others' interests and concerns are most important; consequences in light of everyone's interests and concerns are considered; uncritically doing what others are doing is the criteria; values are weighed and related to ideals, religious, and/or moral concepts.

When the value orientation of a person is known, an understanding of the basis for disagreement is possible. Assumptions or perspectives are realized. Open-mindedness toward those with different orientations is encouraged. One may still not accept other's values for himself. He may agree to disagree. An adage is "Don't judge an Indian until you have walked in his moccasins."

There is a wide range of concerns and perspectives which guide behavior in the society. In a mobile society with mass communication there is blurring of these orientations. Dalke described in 1958 various orientations he saw operating in the society.

A religious oriented person has as his ultimate goal to do God's will. Personal development for service to others is important. Wealth and property are not valued highly. Simplicity in living to free the spirit for more important ideals is stressed. In contrast, the nativist oriented person lives for the honor of the national culture and/or state. Wealth and property are valued as an expression of and in service of the nation-state. In the market value orientation, a person's goals are goods, profits, prestige and wealth. His attitude toward wealth is that its accumulation is the supreme good and a symbol of respectability. The common man oriented person works for the dignity of individuals or workers. A comfortable level of living is regarded as desirable but unlimited accumulation as unsocial. In the humanist value orientation, a person holds as ultimate ends knowledge, creativity and man as the measure of things. Wealth and property are regarded as means for personal and community development. Things are needed as instruments for action and the aesthetic life. Using the scientific method of solving problems is a virtue. Sensitivity to others is important (Dalke, 1958).

Families at different periods of their existence exhibit differences in their values, goals, responsibilities, resources, roles, problems and composition. These different periods have led to the development of the concept of a family life cycle. The family life cycle is an expression that refers to the life span of a family from marriage to the death of one or both spouses. It is based on the assumption that families having children pass through a series of stages characterized by behaviors and responsibilities related to the bearing of children, rearing and launching them into adult pursuits. The concept tries to explain families' behavior and activities that are expected to occur during these stages or phases. The various stages may represent life styles or characteristic patterns of living at particular intervals and are influenced by family composition, age, interest, and activity. Families may or may not follow these patterns exactly; divergence can occur in the individual family.

One description is related to the age of the oldest dependent child. The stages are (1) adjustment, (2) accumulation, (3) grade school, (4) high school, (5) college stage and/or launching, (6) recovery or rediscovery, and (7) retirement. Stages can be adapted for classifying families with more than two generations as (8) extended or (9) families with multiple marriages.

Contributions and Limitations of Factors

Factors chosen for studying life styles have contributions and limitations and, therefore, should be chosen for the specific purpose of the study. The typological study as a method also has limitations.

The scientific value of the constructed type is to (1) identify, (2) simplify, (3) guide initial selection of data, (4) interpret particular situations, (5) have a general standard by which a concrete occurrence is comprehended, (6) generalize concepts by means of which one can extract its empirical versions from different cultural contexts, (7) classify significant, although not merely a class, and thereby differentiate phenomena and set the stage for prediction, and (8) have a point of reference for analysis of social order as it serves as basis for comparison and potential measurement of concrete occurrences. The type does not necessarily refer to the most common form of a phenomena, but usually to the most significantly representative form (McKinney, 1966). The limitations to typological study include the problems that (1) it is still largely in the prequantitative state, (2) not everyone fits into the typological order, (3) the model types represent different portions of families (consumers), (4) lines are not clear cut between types, and (5) some families are mobile and have different characteristics than those that remain in the type.

Values as a factor for analysis of life styles indicate priorities as explanation of consumer behavior. The problem with using this factor is that explicit values may not be the same as a person's implicit values, persons may select from various value systems to have a unique one, and one person's values cannot be used to represent the family's. Basically, most people probably have the same values. The concept of different values refers to differences in rank order, interpretation, degree of emphasis and manifestation. There are several types of values such as cultural, economic, and social. The type will have to be chosen for the given research purpose. Identification of another's values is difficult because an activity may represent several different values and various values for different people. Another problem with value analysis is that frequently persons are described as security-comfort oriented as opposed to those described as achievement-risk oriented. However, the secure person is the one more able to take risks.

Goods already obtained as indicators of life style or level of living may be easier to identify and quantify than other factors. They do reflect past values and income. Therein lies the limitation. Goods are a result of past decision, irrevocable decisions, and what is available in the market. Analyzing a budget of time or money has limitation in that value strength is not proportionately related to expenditures for the same reasons plus the costs of certain goods and activities vary due to their nature.

The family life cycle stages indicate the probable life style of the given consumers in the different stages. However, values probably are the same throughout the stages although the goods and manifestations are different. Also, income and net worth size are generally related to the stages.

Socio-economic class identified by using occupation and education is useful as a typology but it has the limitation of identifying those that are mobile and blurred class lines. Also, the classification suitable for the unemployed or retired is difficult to determine. The uniformity of goods obtained in our society conceals class membership. The social ideal which stresses equality combined with mass production, the nature of the market system and mobility is responsible for the uniformity of goods obtained irrespective of group membership. The "standard package" (Riesman & Roseborough, 1960) refers to the necessary package of economic goods families perceive as necessary for the American style of life. Attitudes toward the "standard package" vary by class as, also, does the manner in which it is provided. The upper class may level down to accept the package with less strain. If success is achieved, it is played-down when it exceeds the friends'. The lower class achieves it with more effort and runs the risk of losing it with poor management or loss of income. Variations in the package are of subtle quality differences.

Resources and motivations vary by socio-economic class. Concepts vary as do the attitudes toward goods. The future to the middle class may mean a rosy horizon whereas to the lower class it is non-existent so that each moment must be lived fully. Problems vary, also, and must be understood in order to identify means for improving levels of living. Knowledge of alternative ways of living is gained by exploring the differences in life styles.

Consumer behavior can be analyzed by examining motivations for consumption. Foote suggested that the most constructive developments in the selection of predictors of household decision-making have come from the study of intervening variables such as motives, attitudes, authority and values (Foote, 1961).

Results of a study of the relationship between investment behavior and social class pointed out the limitations of using income alone as a predictor. With income held constant, white collar classes invested more in education, medical care and insurance than blue collar families. Unskilled workers invested in consumer durables. Brown found that when family resources are limited, social class is more important than income in allocation of investment funds. She concluded that white collar families are more concerned with the future than were families of blue collar workers (Brown, 1969).

The financial problems families have experienced were found significantly related to economic well-being, income, and perceived adequacy of income in research analyzing the data from NC90 project. However, some problems were experienced at all income levels. The higher socio-economic families had certain problems more frequently than other families (Unpublished Theses, 1971-72). It was concluded that problems are a function of priorities, management, and income. Also, problems to families at various income and socio-economic levels have different meanings.

The manner of using goods reflects the standards with which one lives. Standards may be used as basis for comparisons or the criteria for use of resources. Standards are a measure of quality and/or quantity which reflect the reconciliation of resources with demands upon them. The standard of living or goods that make up the standard are those for which the consumer feels deprived when he does not have them. In this sense, standards are motivations for behavior and reflect aspirations and attitude toward goods obtained. The standards are personally identified in this concept rather than being an outside criteria for evaluating life style.

Theoretical Model

The theoretical model attempted to synthesize the concepts suggested in the hypothesized components of life style and descriptions of social classes, by various writers. Activities, attitudes and aspirations are expressed in the typological presentation. The purpose for this typological model is to gain insight into consumer behavior and have points of departure for other sub-level types of life style such as those consumers which might be described as mobile. The theoretical model was found to be more useful in analyzing case studies for consumer behavior in life styles and less ambiguous than some other socio-economic divisions.

Indicators for Life Style Analysis

The indicators selected to analyze case studies were needs that dominate, problems to be resolved, attitude toward decision-making, standards that dictate, commodities and services already obtained, attitude toward these commodities and services, values that guide and motivate, philosophy that integrates, sources of satisfactions, bases for security, income that limits, resources available that are used or not used, socio-economic class represented, and influences that affect the life style.

Selection of Indicators

These indicators were selected on the basis: (1) were the commonalities found in the various sources describing life styles; (2) were more specific than the components in the hypothesis; (3) could be operationally defined in specific terms; (4) applied to case studies and found could be mutually exclusive by students who had studied the factors; (5) and involved the dynamics of consumer behavior. These indicators describe in researchable terms the components hypothesized to make up life style: commodities and services used, the manner in which they are used, satisfactions obtained, pattern of use of all resources available, and attitude about the level of living obtained. The indicators were tested in analysis of case studies after students had studied typologies and the theoretical model. Responses could be tabulated and objectively compared. The indicators were useful in analyzing consumer behavior, components that could change consumer behavior, and differences in life styles. Further research is needed to define the quantification of sub-levels of the indicators for statistical analysis. Various sublevels could be chosen depending on the unique purposes of the research.

Life Styles of Social Classes

	Lower	Working	Middle	Upper-Middle	Upper
Goods	Low quality clothes, auto, furnishings	Auto furnishings gadgets	House (owned) Appliances	Latest style commodities and services	Quantity and quality of commodities and services
Manner in using money and goods	Subsistence Week to week terms	Comfortable in living Spent & enjoyed	Carefully handled Respectable and good Deferred gratification	Sophisticated Quality increased with home skills	Conspicuous and inconspicuous consumption Individualistic
Sources of satisfactions and security	Impulsive pleasures Having goods	Family and consumption pleasures Goods Grateful love of children	Children Job Religion	Expressive and aesthetic experiences Cooperation with others Successful career	Reverence for past culture and arts Family Wealth and prestige
Attitudes toward decision-making	Apathetic Not problem-solving	Cautious Thrifty Habitual Security-oriented	Rational Willing to take risks In control	Based on facts & figures Plans for future accomplishments Active mastery	Traditional
Attitude toward progress	Pessimistic Present oriented Feel inadequate to deal with unknown	Maintenance of status quo	Optimistic Future oriented	Actual participation in community & affairs of the world Manipulation of the environment	Value man more than accomplishment Service to others
Use of leisure time	Friends Relatives	Television Activities at home	Magazines Unions, Lodges fraternal orders	Bridge, cocktails Museums, symphonies, yacht & college clubs	Travels
Means for achievement	Consumption goods Irregular employment	Semi-skilled and operative labor Self-reliance	Occupation Education Personal labor to keep homes	Independent in business & professions Personal skills Education	Inheritance of wealth and position or Wealth by good fortune & own efforts

Selection of Case Studies

Case studies to be used for testing the indicators selected for analysis of life styles were collected from over 100 students. The students were asked to write about families they knew. Then they studied concepts involved in life style components and given typologies. This was followed by students applying the list of indicators in analyzing other case studies than their own case study. By comparing the analysis and the case study, four persons on a panel of judges selected the six most representative of socio-economic classes described by patterns of resource use. These final six case studies were further tested by over 100 other students who had not written any case studies but had studied factors involved in life style. The students were able to use the indicators selected for analyzing the life styles in depth. The responses were generally uniform among students and could be tabulated easily, quickly and objectively. This finding may support the validity and reliability of the factors as research indicators.

Conclusions

Individuals and families demonstrate different styles in performing the functions of maintaining their social and economic systems. The styles of living are a result of ideals and norms continued by the economic and social system as a whole. They reflect resource use based on values both habitually developed and consciously acquired. Limitations and opportunities for development of life styles are provided by governmental regulations and provisions. Life styles are controlled, further, by production and consumption activities. They are affected largely by habit but new events in the society alter life styles. Moreover, life styles are a determinant of resource use.

Research on how life styles are developed or changed indicates how consumer behavior is changed. Habit as a result of established ideals and norms; routine consumption; governmental regulations and provisions; income from capital; durable goods and production; and association; and advertising are influences that can be studied.

Development of a life style begins with identification of its components. Life styles are identified, consciously or unconsciously, when: decisions are made, attitudes toward goods are noted; goods are obtained; and future wants expressed. The standard of living is the desired life style.

Influences on life style must be recognized. Aspirations and goals shape the life style. A gap between the standard of living and the level of living provides motivation. If the gap cannot be reduced, certain reactions may follow: aspirations may be lowered or changed; evaluation of methods of reaching them may result in more efficient use of resources, or frustration or rebellion may occur. The individual's personality disposition and the situation determine which reaction may occur.

After knowledge of various value orientations and the means of achieving them is gained, a new life style may be further defined. Then a focus on ultimate goals should be made for a style to be developed. These may be values, life commitments, or a purposeful philosophy. The new life style can be practiced by using resources and goods in day-to-day activities. Attitudes can be adjusted for appropriateness to the ultimate purpose.

An integrated value system is one's philosophy of life. A workable and consistent philosophy enables resources to be used for maximum satisfaction. Predictable behavior is possible.

Since society is complex and culture allows much variety in patterns of behavior, typological structures would seem to help people to understand choices available. Identifying relevant indicators may help avoid oversimplification concerning the nature of life styles in consumer behavior.

Terms have to be defined and the scope of research delineated. Several components in life style and indicators have been suggested on which research in consumer behavior can concentrate.

Footnote

1. Assistant Professor of Home Management and Family Economics, School of Home Economics, Purdue University.

References

- Baker, Georgianne. Abstract of Research Report: Decision-Making and Value Orientation of Guatemalan Peasant Families. Department of Home Economics, Arizona State University, Tempe, AHEA Conference, June, 1972.
- Beyer, Glenn H. Housing and Society. MacMillan Company, New York, 1965.
- Brown, Kathleen H. Social Class as an Independent Variable in Family Economics Research. Journal of Consumer Affairs, 1969, 3, 127-136.
- Dalke, Otto H. Values in Culture and Classroom. Harper and Brothers, New York, 1958.
- Demby, Emanuel H. Over-the-Counter Life Style. Psychology Today, April, 1972, 75-110.
- Engel, James F. Marketing View of Consumer Behavior. Paper presented at American Council on Consumer Interests, 1972.
- Foote, Nelson, (Ed.) Household Decision-Making, New York: New York University Press, 1961, p. 26.
- Maslow, A. H. Motivation and Personality. Harber & Brother, New York, 1954, 107.
- McKinney, John. Constructive Typology and Social Theory. Appleton-Century-Crofts. 1966, 14-18.
- Riesman, David & Howard Roseborough. Careers and Consumer Behavior. In Norman Bell and Ezra Voyles, The Family. The Free Press, Glencoe, Ill. 1960.
- Schlater, Jean D. Investigating Values Underlying Family Decisions. Research bulletin 23, Agricultural Experiment Station, Michigan State University, 1969.
- Unpublished theses. Purdue University, 1971-72.

ON THE PREDICTIVE ACCURACY OF SUBJECTIVE PURCHASE PROBABILITIES

Donald H. Granbois and John O. Summers¹
Indiana University

Empirical study of the household decision process for major purchases has relevance for at least three theoretical topics in consumer behavior. Purchases of items such as automobiles, major appliances and vacations are infrequent and costly, and they often involve joint consumption or use behavior. Therefore, these purchases are likely to involve more than one family member (Davis, 1970, p. 170). Such decisions are thus ideal for investigating family decision-making role structure. For similar reasons, the process of deliberation and searching out information on available alternatives and their characteristics is likely to be more complex and fully articulated than is true for nondurables (Katona & Mueller, 1954), and these decisions therefore enable study of the character of these phases of the decision process. Finally, the concept of intention (or problem recognition) and its relationship with the outcome of the decision process is ideally studied in the context of major purchases, since the length of time separating intention and outcome is apt to be long enough to permit relatively independent estimates of each (Howard & Sheth, 1969, p. 133). The study we will report here is most relevant for this third theoretical topic and considers the degree to which purchase intentions measured with a subjective probabilities instrument can predict purchase behavior.

From a policy-making perspective, both producers and governmental policy makers are concerned with such operational problems as short-run forecasts of demand for total durable goods. The magnitude and volatility of this demand has important implications for both marketing decisions and economic policy. Forecasting models typically include purchase intentions and fulfillment rates as variables. Conceptually, period-to-period changes in durable goods demand result from non-offsetting changes in the proportion of households having purchase intentions, the rate of intentions fulfillment, and the purchase rate among "non-intenders" (Juster, 1966). Empirical study of the decision process and the role of intentions is relevant here, too, since discovery of determinants and patterns of interrelationships among these three components may suggest significant improvements in forecasting models based upon data regularly collected in purchase intentions surveys (Granbois, 1971a). We think our findings may have relevance for such model building.

Strategies for Improving Forecasts

Purchase intentions data have been integrated into forecasts since the Survey Research Center at the University of Michigan first started collecting them in 1945. However, these data have been considerably less predictive of actual behavior than desired. Longitudinal studies in which respondents' later behavior has been measured have revealed that the traditional survey data are poor predictors of short-run changes in demand because fulfillment rates are low, purchase rates among "non-intenders" are high, and yet these rates are not highly correlated over time so they cannot be systematically adjusted by the analyst (Juster, 1966).

At least five strategies seem feasible for improving the accuracy of forecasts based upon intentions data:

- (1) Improve the initial measure of intentions, so as to reduce the magnitude and variability of the "unplanned" purchase rate and to increase the magnitude and reduce the variability of the fulfillment rate. In the last few years, simple intentions questions have been replaced by measures of subjective purchase probabilities that elicit responses in terms of "Chances out of 10" (or 100) of purchase during three-, six- or twelve-month future periods. These probabilistic questions provide significantly better predictions of purchase rates than do intentions questions (Juster, 1966; McNeil & Stoterau, 1968; Granbois & Willett, 1968; Clawson, 1971).
- (2) Study the relationships between intentions, fulfillment and unplanned purchase rates and the contingencies (unexpected expenses or income, changes in prices, unemployment, etc.) with which they may vary. At least one attempt has been made to prepare separate forecasts of these contingencies and to modify the forecasted rates accordingly (Kosobud & Morgan, 1964).
- (3) Study the consumer search process intervening between intentions and outcome to find the effects of new information on consumer intentions fulfillment and unplanned purchase rates (Pratt, 1965).
- (4) Disaggregate so that separate forecasts are made for categories of products for which relationships between intentions, fulfillment and unplanned purchase rates differ, or for categories of products for which these rates respond differently to **exogenous variables**, such as price levels and income change. Categories may involve different generic products (refrigerators, dishwashers, etc.) for which such dimensions as length of planning period may vary (Pratt, 1968); different kinds of purchase circumstance, such as first-time acquisition or replacement (Heald, no date); or different degrees of urgency or priority of acquisition (McFall, 1969). Another basis for disaggregation might be demographic segments, such as income groups. (Namais, 1960).
- (5) Study patterns of relationships between intentions, preferences and outcome for both husbands and wives (and perhaps older children) since certain family members may be better predictors of household behavior than others, and since degree of conflict or difference may be a key determinant of fulfillment rates. (Granbois, 1971b). Comparisons of husbands' and wives' responses have revealed discrepancies in studies of family participation in major purchase decisions (Granbois & Willett, 1970; Davis, 1970) and in measures of major purchase plans and plan fulfillment. (Wolgast, 1958).

The design of our study was influenced by three of these strategies. We collected subjective purchase probabilities for future major purchases (Strategy #1) from husbands and wives independently and from the same couples jointly (Strategy #5), and later determined the actual purchase behavior of these families. Our sample, though small, permitted some disaggregation by product category and demographic segment (Strategy #4). We did not study intervening search behavior (Strategy #3) or the effects of contingencies (Strategy #2), although our method could easily be modified to include these factors.

Methodology

A group of 77 married couples participated in a behavioral laboratory study of decision processes for major expenditures,² in which one instrument asked for major purchases (\$100 or more) planned during the coming year. A subjective probability in the form of an eleven-point "Chances out of 10" scale, estimated cost, and expected month of purchase were required for each item listed. The

instrument was first administered to each husband and wife separately; each couple then repeated the exercise, discussing their responses and completing a joint form representing their consensus. Responses were not restricted to durable goods plans since these were thought to interact with other major expenditures (vacations, investments, etc.).

A mail questionnaire one year later determined actual purchase behavior. The follow-up questionnaire asked for purchase details (or reasons for not purchasing) for each purchase plan listed in both the individual and joint phases of the laboratory study, and for any other major expenditures during the period. Changes in household circumstances, income, working status of husband and wife, and other expenses were also measured.

Analysis and Results

Our results involve three types of evaluation of subjective purchase probability responses:

- (1) Analysis of the comparative predictive accuracy of aggregated husbands' responses, aggregated wives' responses, and aggregated responses given by husbands and wives jointly.
- (2) Evaluation of the predictive accuracy of response for individual purchase item categories, such as automobiles, vacations and travel, etc.
- (3) Evaluation of several possible determinants of the variability in predictive accuracy of response including age, income and wife's work status.

Analysis of Aggregate Responses

The 77 couples produced more plans responding jointly than either husbands or wives responding singly; expected purchase rates (mean subjective purchase probabilities) for each set were quite similar; and purchase rates varied somewhat over the three response sets (see Table 1).

Subjective probabilities were somewhat concentrated on even numbers, probably because verbal cues were attached to these numbers on the response form. With a few exceptions, purchase rates tended to increase uniformly from low to high probabilities, as expected. In each set, expected rates considerably exceeded actual purchase rates, a finding also reported by Clawson (1971) for a quite different set of items, but contrasting with Juster's (1966) results. On the basis of this simple aggregate analysis, there is little basis for favoring husbands', wives' or joint subjective probabilities as more accurate predictors of purchase rates, although the discrepancy between actual and expected purchase rates was greatest for wives and least for joint responses. Because these averages were aggregated over several item categories and over respondents varying in such demographic characteristics as age and income, they may have disguised several important variations in incidence of plans and in the relationship between purchase probability and outcome.

Variations by Item Category

Purchase plans were classified into nine categories and incidence of plans, expected purchase rates (mean subjective purchase probabilities) and purchase rates were computed for each category. Table 2 summarizes these data for each of the three response sets.

Wives listed many more plans for furniture and carpeting than did husbands,

Table 1

Comparison of Actual and Expected Purchase Rates by Level of Subjective Probability for Husbands', Wives' and Joint Response

Subjective Probability	Joint Responses			Husbands' Responses			Wives' Responses		
	Number of Plans	Actual Purchase Rate	Actual Minus Expected	Number of Plans	Actual Purchase Rate	Actual Minus Expected	Number of Plans	Actual Purchase Rate	Actual Minus Expected
0	6	0.0%	0.0%	7	0.0%	0.0%	7	0.0%	0.0%
1	9	22.2	12.2	4	25.0	15.0	1	0.0	-10.0
2	19	10.5	-9.5	19	21.1	1.1	19	15.8	-4.2
3	12	25.0	-5.0	9	11.1	-18.9	6	0.0	-30.0
4	59	37.3	-2.3	41	34.1	-5.9	57	29.8	-10.2
5	22	45.5	-4.5	16	43.8	-6.2	10	40.0	-10.0
6	48	64.6	-4.6	46	52.2	-7.8	42	52.4	-7.6
7	16	56.3	-13.7	14	42.9	-27.1	6	83.3	13.3
8	68	60.3	-19.7	60	58.3	-21.7	60	58.3	-21.7
9	31	64.5	-34.5	23	78.3	-11.7	17	76.5	-13.5
10	106	84.9	-15.1	88	88.6	-11.4	80	78.7	-21.3
Total	396			327			305		
Means		58.1	-10.6		57.5	-11.4		53.1	-14.5

Table 2
 Comparison of Actual and Expected Purchase Rates by Item Category for
 Husbands', Wives' and Joint Responses

Item Category	Joint Responses			Husbands' Responses			Wives' Responses		
	Number of Plans	Expected Purchase Rate	Actual Minus Expected	Number of Plans	Expected Purchase Rate	Actual Minus Expected	Number of Plans	Expected Purchase Rate	Actual Minus Expected
Furniture and Carpeting	81	59.5%	-12.6%	52	61.4%	- 9.5%	91	59.2%	-18.5%
Travel and Vacation	63	77.0	- 7.2	55	78.7	- 7.8	44	82.5	- 5.2
Home Additions, Remodeling and Repair	54	64.0	-19.6	46	66.5	-23.0	41	61.5	-20.0
Auto Purchase and Repair	51	68.9	9.5	45	75.9	4.1	36	72.5	- 5.3
Miscellaneous	51	87.5	-15.0	36	84.5	- 6.7	28	84.6	-20.3
Appliances	42	57.0	-18.9	43	55.5	-25.3	36	57.8	-27.2
Home Entertainment Equipment	33	54.5	-12.1	30	54.0	-24.0	19	64.7	-17.3
Clothing	12	87.5	4.2	12	84.1	- .8	8	81.0	6.5
Services and Investment	9	89.0	-22.3	8	82.5	- 7.5	2	100.0	-50.0
Total	396	68.7	-10.6	327	68.9	-11.4	305	67.6	-14.5

and husbands' plans were more evenly distributed among categories than were wives' responses, but the rank order of plans by category was quite similar for all three sets of responses. Spearman rank correlations were .979 comparing husbands' and wives', .954 comparing husbands' and joint, and .962 comparing wives' and joint responses. Clothing appears as a category because several respondents anticipated "clustered" purchases expected to cost \$100 or more. In almost every category, more plans were generated by joint responses than by husbands or wives responding individually, partly because "new" plans were created during discussion and partly because joint responses represented a merger of items originally listed by just one spouse but agreed upon by both during discussion.

Expected purchase rates by purchase category exhibit two striking characteristics: Substantial variation appears across categories (54.5% to 87.5% for joint responses, for example) and very close correspondence across all three sets of responses. Spearman rank correlations were .817 comparing husbands' and wives', .944 comparing husbands' and joint, and .870 comparing wives' and joint responses. Highest mean probabilities were assigned to clothing, services and investments, and miscellaneous items (an extremely heterogeneous category including cameras, art, watches, sporting goods and recreational equipment, etc.). Lowest expected purchase rates were given to home entertainment equipment, appliances, and furniture and carpeting plans. High probabilities appeared to be associated with items purchased on a seasonal or time-specific basis ("clustered" clothing purchases and travel and vacations) and with items for which a specific item was contemplated (services and investments and miscellaneous). Items for which low probabilities were assigned were either in categories where an uncertain or discretionary need for replacement triggered the plan, where the item was purely discretionary (and therefore both postponable and likely to compete with other discretionary purchases) or where the purchase was contingent on finding the "right" style, color, or price, as in furniture and carpeting.

The difference between actual and expected purchase rates can be regarded as a measure of the predictive accuracy of subjective purchase probabilities. Except for auto and clothing plans, expected rates always exceeded purchase rates, with considerable variation in the magnitude of these differences across purchase categories. Inspection reveals a slight systematic relationship between the magnitude of the expected purchase rate and the difference between actual and expected rates, in that higher discrepancies tended to be associated with lower expected rates. Variations in the difference measure are perhaps also explained by the characteristics of each purchase category involved. Subjective probabilities predicted auto, travel and clothing purchase rates best.³ The time-specific nature of clothing and travel plans may have made them more predictable, and the fact that most respondents had made several auto purchases in the past may have improved the accuracy of these probabilities. Home addition and remodeling plans are likely to involve many uncertainties with respect to costs, materials, design and availability of craftsmen. Thus, unexpected difficulties and delays may have depressed the actual purchase rates. Appliance and home entertainment equipment plans, both low in expected rates, exhibited large discrepancies between actual and expected rates, perhaps because of the uncertainty associated with replacement need and the almost totally discretionary nature of many items in these categories.

The general pattern of differences between actual and expected rates was quite similar across all three sets of responses, suggesting that the pre-

dictive accuracy of subjective probabilities varies much more across purchase categories than it does among the three respondent types.

Effects of Husband's Age, Income and Wife's Work Status

In contrast to the analysis of the effects of purchase category, few systematic relationships appeared when incidence of plans, expected purchase rate, and predictive accuracy were computed for husband's age, husband's income and wife's work status categories. Table 3 summarizes these findings. A slight tendency for older men to assign higher subjective probabilities to their purchase plans is evident, but the small cell sizes for respondents over 50 and the relatively high probabilities for young men (under 30) suggest caution in generalizing this tendency. Discounting the two lowest income groups because of small cell size, there is a somewhat stronger positive relationship between expected purchase rates and income for joint and husbands' responses. Since wives' responses do not exhibit this relationship, however, this finding should also be interpreted with caution. No clear pattern of relationship between predictive accuracy and either husband's age or husband's income appears. Wife's work status seems unrelated to any of the three dependent variables and, somewhat surprisingly, number of plans per respondent was not related to any of the three demographic variables. With a few exceptions, patterns of response across joint, husbands' and wives' responses were quite similar. Thus, as in the aggregate and purchase category analyses, respondent choice seems to make no difference in the investigation of demographic groupings.

Conclusions

Unlike earlier studies, the design here did not limit respondents to a predetermined list of possible purchases, but rather, they originated their own purchase plans in a kind of unaided response environment. Perhaps for this reason, more plans were generated by husbands and wives interacting than when each spouse listed plans independently. Should the method be used specifically for generating estimates of the magnitude of short-run changes in durable goods demand, it is likely that better estimates could be made using joint response data than responses from either husbands or wives. This conclusion should be tested using a predetermined list of durables instead of the unaided technique used here.

Mean subjective purchase probabilities exceeded actual purchase rates and, although actual purchase rates were positively related to the expected purchase rates, enough variation in the predictive accuracy of subjective probabilities was found to encourage the pursuit of strategies for improving their performance.

On the basis of the findings reported here, disaggregating by purchase category and (to a lesser extent) by demographic characteristics seems to be a more promising strategy for improving predictive accuracy than using husbands and wives responding jointly as respondents in preference to using husbands or wives alone. Research on the comparative responses of husbands and wives should continue, however, since respondent couples in the present study were somewhat homogeneous with respect to education and social class, and the behavioral laboratory environment may have resulted in somewhat different responses than would have occurred in an at-home survey setting. Greater differences between husbands and wives, as well as greater variations among demographic segments, might therefore be discovered in future research.

Table 3

Comparison of Actual and Expected Purchase Rates by Husband's Age, Husband's Income, and Wife's Work Status for Husbands', Wives' and Joint Responses

Household Characteristic	Number of Respondents	Joint Responses			Husbands' Responses			Wives' Responses		
		Plans/ Resp.	Expected Purchase Rate	Actual Minus Expected	Plans/ Resp.	Expected Purchase Rate	Actual Minus Expected	Plans/ Resp.	Expected Purchase Rate	Actual Minus Expected
Husband's Age	14	4.7	65.9%	- 3.8%	3.6	69.6%	-12.7%	3.9	67.6%	-13.9%
	26	4.8	69.8	-14.6	4.0	70.5	-20.0	3.6	70.0	-16.8
	17	5.9	64.4	- 6.0	4.8	63.1	- 3.8	4.4	65.2	- 6.5
	9	5.4	71.2	-14.1	4.4	74.8	-14.8	4.2	66.0	-16.0
	4	5.2	81.9	-10.5	5.2	81.9	- 5.7	4.2	74.1	- 9.4
Husband's Income	5	4.4	81.4	-13.2	3.2	88.7	-19.9	3.2	73.7	-17.4
	2	11.0	62.3	-21.4	8.5	58.8	-17.6	4.0	63.7	-26.2
	14	3.6	55.1	-10.0	2.8	51.0	-10.0	3.0	55.5	-10.3
	20	5.4	67.9	-12.8	4.1	68.9	-12.8	4.2	66.6	-18.4
	11	5.6	67.3	- 4.4	5.0	67.3	-14.6	3.9	73.5	-15.4
	6	4.3	65.0	.4	4.7	57.5	- .4	3.5	76.2	-14.3
	17	5.6	76.9	-11.3	4.9	79.4	- 8.3	4.8	72.3	-11.8
Wife's Work Status	32	5.7	71.3	- 7.2	4.8	65.9	- .5	4.2	68.8	-12.9
	44	4.6	67.4	-12.9	3.7	68.8	-16.7	3.8	66.8	-15.6

Attention should be directed to exploration of additional bases for disaggregation not touched on here. In particular, replacement versus first-time acquisition plans and the extent of previous purchase histories may affect the predictive accuracy of subjective probabilities. Finally, the strategies not examined here--studying the consumer search process and taking into account the role of contingencies--deserve investigation.

Footnotes

1. Donald H. Granbois is Professor of Marketing, John O. Summers is Associate Professor of Marketing, Indiana University.
2. This was part of a joint study conducted by Donald Granbois and Douglas Longman. Responses to three other instruments used and detail on methodological findings are found in Longman (1970).
3. Clothing and service and investment data are based on very small cell sizes, and not much significance should be attributed to findings for these categories.

References

- Clawson, C. J. How Useful are 90-day Purchase Probabilities? Journal of Marketing, 1971, 35, 43-47.
- Davis, H. L. Dimensions of Marital Roles in Consumer Decision Making. Journal of Marketing Research, 1970, 7, 168-177.
- Granbois, D. H. Decision Processes for Major Durable Goods. In G. Fisk (ed.), New Essays in Marketing Theory. Boston: Allyn and Bacon, 1971 a, pp. 172-205.
- Granbois, D. H. A Multi-level Approach to Family Role Structure Research. In D. M. Gardner (ed.), Proceedings of the 2nd Annual Conference of the Association for Consumer Research, 1971 b, pp. 99-107.
- Granbois, D. H. & Willet, R. P. An Empirical Test of Probabilistic Intentions and Preference Models for Consumer Durables Purchasing. In R. L. King (ed.), Marketing and the New Science of Planning. Chicago: American Marketing Association, 1968, pp. 401-408.
- Granbois, D. H. & Willett, R. P. Equivalence of Family Role Measures Based on Husband and Wife Data. Journal of Marriage and the Family, 1970, 32, 68-72.
- Heald, G. The Use of Consumer Attitude and Buying Intentions Data as an Aid to Government Policy Making. (Unpublished manuscript).
- Howard, J. A. & Sheth, J. N. The Theory of Buyer Behavior. New York: John Wiley, 1969.
- Juster, F. T. Consumer Buying Intentions and Purchase Probability: An Experiment in Survey Design. Journal of the American Statistical Association, 1966, 61, 658-696.
- Katona, G. & Mueller, E. A Study of Purchase Decisions. In L. Clark (ed.), The Dynamics of Consumer Reaction. Consumer Behavior, Volume 1. New York: New York University Press, 1954, pp. 30-87.
- Kosobud, R. F. & Morgan, J. N. Consumer Behavior of Individual Families Over Two and Three Years. Ann Arbor: Survey Research Center, University of Michigan, 1964.
- Longman, D. S. A Laboratory Study of Husband-wife Interaction in Consumption Decision Making. Unpublished doctoral dissertation, Graduate School of Business, Indiana University, 1970.
- McFall, J. Priority Patterns and Consumer Behavior. Journal of Marketing, 1969, 33, 50-55.
- McNeil, J. M. & Stoterau, T. L. The Census Bureau's New Survey of Consumer

- Buying Expectations. 1967 Proceedings of the Business and Economic Statistics Section, American Statistical Association, 97-113.
- Namais, J. Intentions to Purchase Related to Consumer Characteristics. Journal of Marketing, 1960, 25, 32-36.
- Pratt, R. W., Jr. Understanding the Decision Process for Consumer Durable Goods. In P. D. Bennett (ed.), Marketing and Economic Development. Chicago: American Marketing Association, 1965, pp. 244-260.
- Pratt, R. W., Jr. Using Research to Reduce Risk Associated with Marketing New Products. In R. Moyer (ed.), Changing Marketing Systems. Chicago: American Marketing Association, 1967, pp. 98-104.
- Wolgast, E. Do Husbands or Wives make the Purchasing Decisions? Journal of Marketing, 1958, 23, 151-158.

DIFFERENCES BETWEEN INTENDERS AND NONINTENDERS
--A METHODOLOGICAL QUESTION¹

Raymond H. Suh²
California State University, Long Beach

Buying plans and attitudes toward durable goods have received primary emphasis in the intentions data collected from consumers. Most individual durable goods purchased involved a substantial expenditure, are infrequently purchased by any single household, and provide considerable latitude in the timing of their acquisition. These characteristics provide analytical support for the belief that many households plan their purchases of durables well in advance and are able to offer accurate information about their decisions in interview.

The basic idea behind a survey of consumer buying intentions is that consumer purchases, of items such as houses, automobiles, and appliances, are subject to fluctuations that are in some degree independent of the effects of socio-economic variables. Fluctuations in these forms of expenditures are believed to be more accurately predicted by both the changes in socio-economic variables and consumer pessimism and optimism than by socio-economic variables alone (Juster, 1966, 1964).

One of the problems for an intentions survey is its inefficiency of the basic predictors of purchase rates. As a method of data collection, responses are generally classified into several categories such as "definitely will buy," "probably," and "no," for the expression of buying intentions. The usefulness of the survey is then evaluated by relating variations in the fraction of one or more groups of intenders (respondents reporting "definitely" or "probably") to variations in the fraction reporting purchases. It is puzzling in most studies that data represent only the intenders; and there is no survey evidence that bears directly on the predictability of the critically important movements in nonintenders' purchase rates.

Namias' (1960, 1959) studies indicated that consumers who do not plan to buy household durable goods are more consistent in carrying out their no-buying-intentions than those who plan to buy. On the other hand, a large number of consumers, although small in population, who do not plan to buy, change their minds. Juster (1966, 1964) noted that the majority of durable goods purchases are made by households that are generally classified as nonintenders by consumer surveys. As a consequence, the accuracy of purchase predictions based on intentions surveys depends largely on whether or not changes in proportion of intenders can successfully predict changes in the purchase rate of nonintenders being strongly correlated over time. Piskie (1963) has provided similar arguments in his study. However, Piskie argued that researchers should not only have purchase data on their subjects but should also know subjects' buying plans and the degree of certainty of those plans.

It also appears that buying intentions data are collected, in general, for the wife as a sole respondent who is assumed to represent the family's buying plans. Rarely did studies look into the differences between the husband and wife when contemplating buying intentions, and their subsequent effects on the accuracy of the prediction of the durable goods purchases. Some skepticism, therefore, does exist concerning the method of data collection for family decision making. For example, as early as 1955, Ferber said:

The method by which those studies have been carried out by soliciting the opinion of one family member regarding the purchase influence possessed by himself and by other members of the family and the scarcity of any rigorous tests of the validity provides sufficient basis for receiving such studies with skepticism (Ferber, 1955a, p. 225).

By and large, the findings derived from a number of works are conflicting and thus not very convincing. Some results, based on interview results from a single respondent in a household, suggested that the amount of discrepancy in a household is small when aggregate responses for husbands and wives are compared. This is not so when the responses from husbands and wives in the same families are compared (Blood & Wolfe, 1960; Granbois & Willett, 1970; **Hear**, 1963, 1962; Sharp & Mott, 1956; Wallin & Clark, 1958). Other disclosures demonstrated that poor agreement between husbands and wives on the family income, durable goods buying plans, actual purchases of the planned buying, and the frequency of sexual intercourse has shown a poor agreement between the members involved (Davis, 1970; Ferber, 1955a, 1955b; Haberman & Elinson, 1967; Suh, 1972).

It appears that these controversies are due to difficulties in data collection methods. Ferber, in two previous works (1955a, 1955b), argued that the reliability of ratings concerning the relative influence of different family members on purchases obtained by direct questioning of only one of the family members is highly limited. He also indicated that the securing of reliable data, at least in consumer purchase studies, is much more difficult than previously purported. Suh (1972a, 1972b) further indicated that the salient motives for husbands and wives are different for contemplating buying intentions, and more accurate data could be obtained by interviewing respondents during a period varying from three to six months and by securing at least two responses from each family.

Therefore, the objective of this study was to determine the differences between intenders (i.e., respondents whose purchase probabilities on living room furniture were substantially high) and nonintenders in estimating purchase probabilities for living room furniture based on four buying intentions scores, eight attitude scores, and eight demographic variables. The cutoff point of intenders and nonintenders was also to be determined. A discriminant analysis was performed for husbands and wives separately, and responses from each sex are compared to examine possible sex differences.

Methodology

The subjects were members of the Illinois - Berkely Panel of Consumer Decision Processes who were recently married couples with the husband being 30 years of age or less. The sample size was 230 couples, 166 living in Peoria and 64 in Decatur, Illinois, both cities being in the 100,000 - 200,000 population range. The sample members were selected by systematic random selection from a list of such couples married between June 1, 1968, and September 30, 1968.

Most questions on durable goods purchases and buying intentions were asked of panel members since Fall, 1968. Interviews were conducted only by personal interviews and at intervals varying from two to six months.

Subjective purchase probabilities were used as a measure of durable goods buying intentions. The use of subjective probability had certain advantages over other scaling techniques such as a dichotomous (yes or no) form of the responses, or gradational adjectives (i.e., "definitely will," "probably will," "maybe," "probably not," and "definitely not") in obtaining information on buying intentions. Advantages included: first, the additional information throws light on the degree of certainty with which particular purchase plans may be carried out; second, introduction of buying plans data in this form may serve to further improve the ability to anticipate actual purchases (Ferber & Piskie, 1965).

The purchase likelihood scales were eleven-point scales ranging from "absolutely certain (100%)" to "no chance whatsoever (0%)". The scale appeared as follows:

Absolutely certain -- 100%
 90%
 80%
 70%
 60%
 50%
 40%
 30%
 20%
 10%
 No Chance Whatsoever -- 0%

Only living room furniture, of the sixteen durable products studied, was used in the analysis because living room furniture produced the maximum possible sample size, 230. There were eight demographic variables, eight attitudes scores and four buying intentions scores (See Table 1). Because this was a by-product of a broader study (Suh, 1972a), the study did not provide a process from which both factor scores for attitudinal classification and buying intentions could have been derived.

Once subjective purchase probabilities and rank order information on the sixteen durable goods were obtained, the joint distribution of the rank order information in which each family member hoped to acquire was examined to determine the extent of agreement for living room furniture between husbands and wives. Because the subjective purchase likelihood information was a measure of buying intentions, this measure was used to classify subjects into intender and nonintender groups.

Analysis

The first task was to determine the cutoff point of intenders and nonintenders on the probability scales. Using the objective criterion, it appeared that the purchase probability 0.5 (50%) would be a reasonable estimation of the cutoff point. Very little empirical evidence was available on this point; however, several researchers (Ferber & Piskie, 1965; Juster, 1964; Piskie, 1963) supported this argument. The selection of the purchase probability 50% as a cutoff point was reasonable in the sense that if a respondent is forced to answer yes or no in regard to a future purchase, the cutoff point on the probability scale would most likely be at .50. Therefore, when the subjective probability is a measure of buying intentions, all those respondents given a .50 probability or above could be placed in the yes category (i.e., intender group) and vice versa.

Once the cutoff point was selected, three questions were of primary interest in this analysis. First, did the two groups occupy different regions in the one-dimensional discriminant space? Secondly, what were the discriminating variables? Thirdly, how well could we explain the classification of subjects into the intender and nonintender groups?

Discriminant Analysis for husbands: Out of the total 218 subjects (Table 1), 93 husbands were classified as intenders and 125 as nonintenders. The intender group was operationally defined as the group of subjects whose purchase probabilities were equal to or greater than 0.50; and the nonintender group was the group of people whose purchase probabilities were less than 0.50.

Wilks' lambda, the test of the null hypothesis that the group centroids occupy the same positions for the two groups, was found to be significant at the 0.005 level. This significance test was performed by a multivariate F-ratio, 2.2999 with 20 and 197 degrees of freedom. This suggested that the centroids of the two groups occupied different positions in the one-dimensional discriminant space. The group centroid was 0.7821 for intenders and -0.1099 for nonintenders.

Inspection of both the mean vectors and univariate F tests in Table 1 indicated that the variable innovativeness (#13) was the single most significant variable ($P < 0.025$) that discriminated between the intender and nonintender groups for living room furniture buying intentions. It was also shown that nonintenders had a smaller mean vector than intenders. It was observed that husband intenders were generally high risk takers, had a better occupation and were less price conscious in their shopping attitudes. However, a noticeable difference was observed in the family size and level of education of intenders. The intenders' level of education was somewhat higher and their family sizes were smaller than nonintenders.

Examination of the scaled vectors of discriminant weights in Table 1 revealed that innovativeness made the greatest positive contribution to the centroid configuration, and to a lesser degree, education, a social factor, occupation, and bargaining. The highest negative coefficients were associated with low price option, a product durability factor, age, channel environment factor, and family size. The items above closely corresponded to the items that were significant in the univariate F tests. The variables and their discriminant coefficients seemed to form a kind of "innovativeness-price dimension," but the meaning of the dimension was not clear.

A multiple discriminant analysis as a procedure to classify individuals into the intender and nonintender groups, was also performed. The classification results were:

		<u>Classified</u>		
		Intenders	Nonintenders	Total
<u>Actual</u>	Intenders	60	33	93 (42.2%)
	Nonintenders	<u>37</u> 97 (44.5%)	<u>88</u> 121 (55.5%)	<u>125</u> (57.8%) 218

Table 1

Discriminant Functions, Means, Standard Deviations, and F Ratios
for Analysis I: Husband Intenders and Nonintenders

Variables	Discriminant Functions			Intenders**			Nonintenders***			F Ratio df=1,216	P
	Normalized Vectors	Scaled Vectors	Convention- alized Vectors*	Means	S.D.	Means	S.D.	Means	S.D.		
Demographics											
1. Place of residence	0.37	2.46	0.67	1.30	0.46	1.27	0.45	1.27	0.45	0.21	-
2. Type of housing	0.04	0.47	0.51	1.75	0.70	1.69	0.71	1.69	0.71	0.40	-
3. Age	-0.10	-4.32	0.13	22.59	2.52	22.97	3.21	22.97	3.21	0.87	-
4. Education	0.37	5.35	0.91	0.46	1.27	0.27	0.69	0.27	0.69	2.01	.25
5. Wife working	0.10	1.95	0.64	2.00	1.36	1.90	1.33	1.90	1.33	0.32	-
6. Occupation	0.16	4.46	0.84	3.88	1.85	3.48	1.86	3.48	1.86	2.50	.25
7. Family size	-0.21	-2.75	0.26	0.17	0.52	0.38	1.06	0.38	1.06	2.88	.10
8. Income	0.00	0.06	0.48	10.87	3.82	10.84	4.04	10.84	4.04	0.00	-
Attitude Scores											
9. Low risk	-0.06	-1.04	0.40	2.95	1.01	3.09	1.19	3.09	1.19	0.86	-
10. Bargaining	0.25	4.03	0.80	1.63	1.16	1.44	1.02	1.44	1.02	1.75	.25
11. Quality-consciousness	0.06	1.06	0.56	1.41	1.18	1.28	1.11	1.28	1.11	0.67	-
12. Economic shopping	0.22	3.97	0.80	1.87	1.16	1.55	1.29	1.55	1.29	3.50	.10
13. Innovativeness	0.44	6.51	1.00	1.68	1.05	1.35	0.95	1.35	0.95	5.75	.025
14. Extravagant shopping	0.11	1.78	0.62	1.48	1.21	1.41	1.05	1.41	1.05	2.10	.25
15. Price-consciousness	-0.15	-2.52	0.28	1.09	1.23	1.22	1.04	1.22	1.04	0.73	-
16. Low-price option	-0.33	-5.74	0.02	0.99	1.15	1.34	1.18	1.34	1.18	4.69	.05
Buying Intentions											
17. Social factor	0.18	4.78	0.86	4.25	2.02	3.79	1.59	3.79	1.59	3.52	.10
18. Product-durability	-0.32	-5.22	0.06	0.98	1.04	1.31	1.17	1.31	1.17	4.82	.05
19. Channel environment factor	-0.19	-3.36	0.21	0.72	1.07	0.89	1.34	0.89	1.34	1.02	-
20. Pragmatic individual factor	0.06	0.93	0.55	0.75	1.19	0.66	1.05	0.66	1.05	0.36	-

* Conventionalized Vectors + (Scaled Vectors + 6) / 6.51 (innovativeness)

** Sample size is 93.

*** Sample size is 125.

The overall accuracy of classification showed that 144 subjects, or 66%, were correctly placed out of 218. Because the maximum chance criterion was 57.8%, 66% classification was not a good fit when considering the fact that the same subjects were used for classification; hence, overestimating the fit.

The proportional chance criterion in this analysis was 51.2%. However, given the chance criterion of 51.2%, 55.5% were classified in the nonintender group; the outcome should have been .509, or about 51% correctly classified. The classification of individuals for the nonintender group was greater than the chance criterion; although the difference was still small, of the 97 individuals classified as intenders, 60 were correct. This was 61.9% compared with a chance percentage of 42.2%.

Discriminant Analysis For Wives: For the analysis of differences between intenders and nonintenders for wives, 89 wives were classified as intenders and 133 were as nonintenders out of 222 wives. Table 2 provides the results for the intender and nonintender groups.

Wilks' lambda was found not to be significant for these two groups. The multivariate F-ratio was 1.1033 with 20 and 201 degrees of freedom. This means that the centroids of the two groups occupy similar positions (i.e., group centroid for intenders was -2.0611 and -2.5755 for nonintenders) in the one-dimensional discriminant space.

Only one variable, type of housing (#2), was significant at the 0.05 level as was shown by the univariate F test. Examinations of mean vectors for the two groups indicated that the intender group was at the low end of the scale, meaning that more intenders were residing in the duplexes or houses which may not be furnished or partially furnished. Other characteristics of intenders were as follows: They were slightly younger, had smaller family size, earned slightly higher income, were higher risk takers, and were less price-conscious than wife nonintenders.

Inspection of the scaled discriminant vectors of discriminant weights in Table 2 showed that the social factor (#19) made the highest positive contribution to the centroid configuration although the group separation was not significant in a statistical sense. Given the 20 variables in the analysis, the social factor had the greatest mean difference; and it was demonstrated that nonintenders were more concerned with the social acceptance and the conspicuousness of the product in expressing a buying intention. The highest negative coefficients were associated with type of housing (#2), family size (#7), low risk (#10), and economic shopping (#11). The variable and their discriminant coefficients seemed to form a kind of "social-socioeconomic dimension," but again interpretation was difficult because the groups were not different in a statistical sense.

The classification of individuals by the function into groups was performed with the following results:

		<u>Classified</u>		
		Intenders	Nonintenders	Total
<u>Actual</u>	Intenders	56	33	89 (40.1%)
	Nonintenders	57 113 (50.9%)	76 109 (49.1%)	133 (59.9%) 222

Table 2

Discriminant Functions, Means, Standard Deviations, and F. Ratios
for Analysis II: Wife Intenders and Nonintenders

Variables	Discriminant Functions			Intenders**			Nonintenders***			df=1,220	P
	Normalized Vectors	Scaled Vectors	Conventionalized Vectors*	Means	S.D.	F	Means	S.D.	F		
Demographics											
1. Place of residence	0.47	3.13	0.91	1.31	0.47	1.26	0.44	0.70	-	-	
2. Type	-0.60	-6.24	0.07	1.58	0.56	1.80	0.78	5.02	0.05	.05	
3. Age	-0.03	-1.05	0.54	20.71	2.31	20.83	2.74	0.12	-	-	
4. Education	-0.11	-1.48	0.50	0.21	1.04	0.32	0.79	0.69	-	-	
5. Husband working	-0.03	-0.28	0.61	1.18	0.63	1.23	0.72	0.33	-	-	
6. Occupation	-0.07	-1.84	0.47	2.10	1.91	2.21	1.77	0.19	-	-	
7. Family size	-0.36	-4.98	0.18	0.21	0.47	0.35	1.09	1.09	-	-	
8. Income	0.05	2.98	0.90	10.99	3.74	10.80	3.97	0.13	-	-	
Attitude Scores											
9. Bargaining	-0.05	-0.80	0.56	1.71	1.00	1.73	1.00	0.03	-	-	
10. Low risk	-0.34	-4.84	0.19	3.08	0.92	3.30	0.95	2.84	0.10	.10	
11. Economic shopping	-0.30	-4.16	0.26	3.78	0.97	3.98	1.09	1.95	0.25	.25	
12. Innovativeness	-0.02	-0.25	0.61	2.80	0.82	2.79	0.83	0.00	-	-	
13. Quality-consciousness	-0.02	-0.26	0.61	2.49	0.99	2.52	1.04	0.03	-	-	
14. Extravagant shopping	0.02	0.23	0.65	3.35	0.96	3.37	1.02	0.01	-	-	
15. Low-price option	0.14	2.03	0.81	5.19	1.09	5.01	0.95	1.62	0.25	.25	
16. Price-consciousness	-0.05	-0.80	0.56	0.01	1.03	0.07	0.99	0.15	-	-	
Buying Intentions Scores											
17. Pragmatic individual factor	-0.10	-1.63	0.48	-0.09	0.82	0.06	1.25	1.02	-	-	
18. Personal-influence factor	-0.03	-0.68	0.57	3.35	1.72	3.36	1.80	0.00	-	-	
19. Social factor	0.14	4.09	1.00	3.17	2.14	2.75	1.83	2.41	0.25	.25	
20. Situational factor	-0.13	3.68	0.30	2.22	1.71	2.61	2.04	2.11	0.25	.25	

* Conventionalized Vectors = (Scaled Vectors + 7) / 4.09 (social factor)

** Sample size is 89.

*** Sample size is 133.

The proportional chance criterion was 51.96%, and the maximum chance criterion was 59.9%. In this analysis, 132 subjects or 59.5% were correctly classified out of 222. Because the maximum chance criterion was 59.9%, the 59.5% correct classification was a poor fit. Given the proportional chance criterion of 52%, 51% were classified as intenders and 49% as nonintenders. Both were poor classifications. Of the 113 subjects classified as intenders, 56 were correct. This was 51% to be compared with a chance percentage of 40%.

Discussion

The results of discriminant analysis for husbands and wives demonstrated some interesting phenomena. Before comparing the results from the two different sex groups, there are several points worth mentioning about discriminant analysis and the classification procedure. First, close inspection of Table 2, for example, demonstrated that the variables that were significant under the univariate F tests and the scaled vectors that had either the highest positive coefficient or the highest negative coefficient did not necessarily contribute largely to the group centroid configuration. It was rather deceiving. The reason was that in this analysis, scale units of the variates were of several different types; scale units were not the same for all 20 variates. Variables from 9 to 16 (i.e., attitudinal classification scores), for example, could assume the values from 1 to 5, and buying intentions scores (variables 17 through 20) could assume the values from 1 to 7; while present income could take on any value from 1 to 100, or age could assume any values from 1 to 30. The discriminant weights, on the other hand, were independent of the units of measurement and the origin of coordinates of the original variates because the coefficients automatically adjusted themselves to the scales employed. Therefore, the problem occurs when the weights were multiplied by the means, and these products were summed across variates--the sum (centroid) was affected not only by the value of the discriminant vector but also by the scale units of the particular variable.

In view of the above discussion, it could be seen that discrimination on the variable, a social factor, was not really high for the wife groups because the scale units of the variate were small relative to the income variate. Thus, the variable making the highest positive contribution to the centroid configuration was the low-price option, and to a lesser but still significant degree, place of residence, income, and a social factor.

Secondly, the twenty variables used in this analysis contributed a great deal to the centroid configuration for the husband groups. Also the classification of husbands into these groups was fair. However, this was not true for the wife groups. There was no statistical significance in discriminating the intender and nonintender groups for wives, and the classification was poor. It should be noted that multiple discriminant analysis classified only on the variables given in the analysis. Poor discrimination between groups could be due to the fact that the most important variables for a particular group of people were overlooked, and hence were not included in the study. Therefore, it may be necessary in the future to select more relevant variables; and it should be possible to assess costs of misclassification of individuals into the designated groups.

Results from the discriminant analysis demonstrated noticeable differences between the groups in expressing buying intentions. With the lack of actual purchase data, it was difficult to assess as to whose buying intentions more accurately described the family's future buying plans. At the moment these

study results appeared to support the argument that it is necessary to obtain two measures of buying intention, one from the husband and the other from the wife, from each household rather than relying on a single measure of buying intentions. These two separate measures of buying intentions may help to improve the prediction of the actual purchases.

The differences between husbands and wives could be further supported by examining the joint distribution of the rank order of the purchase priorities on living room furniture given by husbands and wives. Table 3 describes the rank order of each of the sixteen products (Suh, 1972a) from the first choice to the fifth choice. Because the number of subjects whose rank order was less than six was negligible, those were combined. Diagonal entries indicate the number of couples who were in agreement in their preferences to purchase a given product.

Table 3

Joint Distribution of the Rank Order of the Purchase Priorities for Living Room Furniture by Husbands and Wives

Living Room Furniture								Total (Wife)
Rank Order	1	2	3	4	5	6-16	0	
1	11	9	1	1	1	1	3	27
2	3	9	3	2	0	3	2	22
3	2	4	9	1	0	0	3	19
4	1	3	3	3	1	2	5	18
5	0	1	1	2	4	4	3	15
6 - 16	1	0	2	2	0	8	8	21
0	3	4	9	3	6	15	68	108
Total (Husband)	21	30	28	14	14	33	92	230

It was apparent for living room furniture that (a) the aggregate of the rank order between husbands and wives was not a perfect match, nor was it a poor one, which was indicated in each of the column and row totals. (b) It was even more apparent that there was very poor agreement within a family given that each of the husbands and wives expressed their buying intentions. For example, of the 27 wives (see row total) and 21 husbands (see column total) who gave first priority to the purchase of living room furniture, only eleven couples from the same families were in agreement. (c) Finally, agreements became poorer as the rank order decreased, except in the case when couples were expressing no-buying-intentions (i.e., zero rank order).

It appeared that the poor agreement shown in the joint distribution had a carry-over effect in the discriminant analysis. As it had been seen in the two different analyses, agreement amongst the variables making contribution to the centroid configuration was poor, and even the subsequent discriminant coefficients for variables were different from one to another. Hence, the study raises serious doubts on consumer surveys which simply rely on a single response from each family. This appears to be more difficult and problematic than the previous findings have reported. Differences in expressing buying intentions between husbands and wives do not simply guarantee the reliability of information obtained from a member of the family.

Footnotes

1. This study is a part of the author's doctoral dissertation at the University of Illinois, Urbana-Champaign, and the author wishes to express his appreciation to Professor Robert Ferber for providing data and assistance.
2. Raymond Suh is Assistant Professor of Marketing at California State University, Long Beach.

References

- Blood, R. O. & D. M. Wolfe. Husbands and Wives: The Dynamics of Married Living, Glenco, Illinois: The Free Press, 1960.
- Byrnes, J. C. An Experiment in the Measurement of Consumer Intentions to Purchase. Journal of American Statistical Association: Business and Economics Statistical Section, 1964, 59, 265-279.
- Davis, H. L. Dimensions of Marital Roles in Consumer Decision Making. Journal of Marketing Research, 1970, 7, 168-177.
- Ferber, R. On the Reliability of Purchase Influence Studies. Journal of Marketing, 1955a, 19, 225-232.
- _____. On the Reliability of Responses Secured in Sample Surveys. Journal of American Statistical Association, 1955b, 50, 788-810
- _____. & R. A. Piskie. Subjective Probabilities and Buying Intentions. Review of Economics and Statistics, 1965, 47, 322-325.
- Granbois, D. H., and R. P. Willett. Equivalence of Family Role Measures Based on Husband and Wife Data. Journal of Marriage and The Family, 1970, 32, 68-72.
- Haberman, P. W. & J. Elinson. Family Income Reported in Surveys: Husbands versus Wives. Journal of Marketing Research, 1967, 4, 191-194.
- Heer, D. M. Husband and Wife Perceptions of Family Power Structure. Marriage and Family Living, 1962, 22, 65-67.
- _____. The Measurement and Bases of Family Power: An Overview. Marriage and Family Living, 1963, 23, 133-139.
- Juster, F. T. Anticipation and Purchases: An Analysis of Consumer Behavior. New Jersey: Princeton University Press for National Bureau of Economic Research, 1964.
- _____. Consumer Buying Intentions and Purchase Probability: An Experiment in Survey Design. Occasional Paper 99. New York: National Bureau of Economic Research, 1966.
- Namias, J. Intentions to Purchase Compared with Actual Purchase of Household Durables. Journal of Marketing, 1959, 24, 26-30.
- _____. Intentions to Purchase Related to Consumer Characteristics. Journal of Marketing, 1960, 25, 32-36.
- Piskie, R. A. A Quantitative Measure of Consumer Buying Intentions. Unpublished Masters Thesis, Department of Economics, University of Illinois, Urbana, Champaign, Illinois, 1963.
- Sharp, H. and P. Mott. Consumer Decisions in the Metropolitan Family. Journal of Marketing, 1956, 21, 149-156.
- Suh, R. H. An Application of Multivariate Statistical Techniques for the Study of Consumer Buying Intentions for Durable Goods. Unpublished Doctoral Dissertation, College of Commerce and Business Administration, University of Illinois, February, 1972.
- _____. Response Patterns of Durable Goods Buying Intentions by Husbands and Wives--A Methodological Question. (In process), 1972.
- Wallin, P. and A. Clark. Cultural Norms and Husbands' and Wives' Reports of Their Marital Partners' Preferred Frequency of Coitus Relative to their Own. Sociometry, 1958, 21, 247-254.

MARKETING APPLICATIONS OF INTENTIONS DATA

C. Joseph Clawson
University of Southern California

It is always a pleasant change of pace for a researcher, just as it is for an actor, to switch from the role of performer to that of critic. How sweet to dish it out with lofty condescension or humorous derision! How bitter to have to take it with open mind and dazzling smile, while hiding clenched teeth and murderous heart!

That is why Bob Pratt's invitation to spend part of my time evaluating the Granbois-Summer paper and the Suh paper at today's meeting was so appealing. However, he also invited me to play the performer role a little too, by offering some thoughts on the marketing applications of intentions data. In this schizoid role, it is my purpose to suggest some criteria that marketing managers can--and some do--apply in evaluating consumer intentions data. These thoughts will be interspersed with illustrations taken from the two papers just presented to you, but are not confined to them.

The marketing manager's intention to use intentions data appears to be directly related to the number of "Yes" answers he can find to five crucial questions in his mind. Let us examine these questions as criteria for marketing applications.

1. Does Our Company Need Intentions Data at All?

This issue really breaks down into two sub-issues. First, "Do we really need consumer sales forecasts of any sort, at this time, in our company?" Second, "If so, should we obtain intentions data in making those forecasts?"

Many companies can honestly answer either or both of these questions with a "No." To save time, we will not inquire when and why this is the right answer, but will assume we are talking today only about marketers who can answer both questions affirmatively.

2. Are the Intentions Data Measured Accurately?

That is, do they measure "true" intentions--at least as of the time they are expressed? This is a technical question, but it is appropriate because marketing applications hinge upon the answer.

For instance, is the sample size large enough to give reasonably accurate measurements of all the variables? Neither of today's studies presents the confidence limits for its sample size. However, this is not too serious an oversight for our purposes. The Granbois-Summers sample of 77 married couples must be large enough because it gives clearcut findings! The Suh sample of 230 couples may be too small because it yields disappointing results.

What about bias? Granbois and Summers have been careful to avoid suggesting products to their respondents, by letting them originate their own lists of items they would consider. However, this systematically excludes just about everyone who would have assigned a 0 or .1 probability of purchasing these same items if they had been presented on a predetermined list. Thus, it inflates

all the predicted and actual purchase rates, so they could not be used for estimating the magnitude of future durable goods demand in any larger population, as the authors carefully admit. Dr. Suh does use a predetermined list. However, he provides data here only on one of the product classes, living room furniture.

As for scaling, both studies utilize an 11-point scale to measure strength of intentions. However, the Suh report first praises the superiority of such a scale over simpler ones, then promptly abandons his 11-point scale and converts it into a 2-point scale, intenders vs. non-intenders. While such a split may be required by multiple discriminant analysis, it blurs the entire picture and no doubt contributes to his disappointing results.

3. Are Valid Predictor Variables Employed?

Professors Granbois and Summers were trying to predict purchases by measuring intentions, whereas Dr. Suh went further upstream, hoping to predict intentions with 20 prior influences. However, we can still ask in both cases if they predicted well, and if they measured all of the truly important factors that influence the outcomes.

The Granbois-Summers project confirms, once again, the familiar positive correlation between strength of intentions and actual purchase rates. They also report what I consider very worthwhile, interesting, but not overwhelming effects arising from age, sex, income, wife's working status, and joint decision-making. However, their substantial discrepancies between expected and actual purchase rates point up the basic difficulty in predicting behavior 12 months ahead. Too many changes take place in the main predictor variable, intentions, as the months slowly roll by.

The Suh study neatly sidesteps the time-lapse problem by measuring 20 predictor variables and one criterion variable--intentions--as of the same time. However, as Dr. Suh sorrowfully but manfully confesses, it may be necessary in the future to select different, more relevant variables. The ones he employed just did not give clearcut, significant overall predictions as between intenders and non-intenders.

One moral for all of us, I believe, is that we should measure changes in purchases, changes in intentions, and changes in causal factors in successive, reasonably short periods of time. Another moral is that we need to discover and use a broader range of genuinely causal variables, such as motives, attitudes, social pressures, in-store promotions, competitors' prices, and de-emphasize the use of the remote influences, such as general personality traits, and "proxy" variables, such as demographics. Of course, if good theory suggests that specific traits and specific demographics are causal, that is another matter. Kassarian has argued this point well in the November 1971 issue of the Journal of Marketing Research.

4. Are the Intentions Data "Actionable"?

The marketing manager also wants to know if he can take any action whatever based on intentions data, or if can only say, "Right on!" or "No way!" as he passively files them away.

One way we can make intentions data more useful is to relate them to the types of consumer choices that a given marketing manager has the best chance

of being able to influence. For instance, I would assume that the marketing manager for General Electric television sets can surely do more to change GE's brand share within the industry than he can to change overall demand for television sets in general.

Yet most published studies concentrate on product class intentions--as in today's papers, the University of Michigan studies, the Bureau of the Census reports, and--yes--in my own article in the Journal of Marketing for September 1971. Why don't we research the best ways to use intentions data to predict brand choices, store choices, feature choices, timing choices, quantity choices, and so on?

Another way to make intentions data more actionable is to select the right time period to cover in the forecast. We should not overlook, for instance, that even very short periods, such as two weeks or three months can be long enough for many companies to take compensatory action. While unable to change the predicted sales level, they can adjust to it by shifting their production rates, inventory size, overtime, local advertising budget, and so on. However, if our management is only concerned with the need for remedial or aggressive actions to change the predicted sales level itself, then of course the planning period should be six months, a year, or five years--whatever period is affected by long-range commitments in new plant and equipment, investment in R & D, and so on.

5. Can Intentions Data Help to Guide Marketing Action?

Guiding action is not the same as alerting action, but intentions data could do both if tied to certain other facts. In fact, the potential pay-off multiplies as they are added.

Of course, even at the bottom rung of this pay-off ladder, where intentions data are completely unadorned with supplementary information, they can at least alert marketing managers to their need to consider possible action. However, the matters of "What," "To whom," "How," and "Whether" are left blank, for them to fill in.

Rising to the second rung, as both of today's studies do, for example, we add demographics like sex, age, area, income, and so on. Such data can actually help managers pinpoint the most fertile market segments, identify problem categories requiring attention, and abandon hopeless situations. This leaves only "What," "How," and "Whether" unanswered.

So to reach the third rung, we can add data on attitudes, motives, confidence, awareness, and other variables that hopefully determine the intentions, as the Suh study commendably tries to do. Here the data become diagnostic. As such, they provide ideas for what needs to be changed, such as price, advertising copy, product features, and so on. The managers, however, must still create their own answers to "How" and "Whether."

The fourth rung on the ladder may be achieved when we include "contingency intentions." Our consumers are asked for their purchase probabilities under various hypothetical conditions, such as a proposed change in product, package, retail outlets, advertising theme--or even imaginary changes in competitors' efforts, personal income, taxes. We still need to study the best ways of getting valid contingency intentions and testing them under controlled conditions. However, if he had them, the marketing manager could be greatly assisted in selecting the right marketing mix for reaching any desired sales level.

Finally, the top level of potential usefulness comes into sight, at least, when the different marketing mixes that we derive from contingency intentions are matched against internal cost figures. Combined, the manager can come closer to picking the optimum mix, assuming anyone cares about profitability or return on investment! There, I think, is a real challenge to the producers and users of intentions data.

In conclusion, I hope that all of us who work with intentions surveys can remember to match them to the needs of our own customers, the marketing executives. We can thank the authors of today's two papers, and more generally, we can be proud of the practical marketing applications of today's intentions data, which are already valuable. However, the untapped potential is far greater.

Buyer Behavior Models and Attribute Models: A Synthesis

Donald R. Lehmann
Columbia University

Introduction

Recently there has been substantial interest in developing models of individual consumer choice. Two of the most popular types of models are large-scale models of buyer behavior of the type proposed by Nicosia (1966) and Howard and Sheth (1969) and attribute models of preference based on the models of Fishbein (1967) and Rosenberg (1956). This paper will begin by summarizing the current state of research in these two areas, and then will suggest how the two types of models might be combined.

Buyer Behavior Models

Buyer behavior models have addressed the question of how a buyer goes about gathering information for making a decision, how he makes a decision, and finally how the decision affects his attitudes and hence future decisions. In other words, they are attempts to describe buyers from "cradle to grave." These models are thus directed at the Herculean task of explaining buyer behavior in every facet.

These buyer behavior models are usually stated in terms of a flowchart. These flowcharts suggest the general direction of flows from one endogenous variable to another. They do not, however, provide operational definitions of the constructs in each box of the flowchart, nor do they in general specify what exogenous variables affect the various endogenous variables. Moreover, they do not specify the mathematical form of the links between variables. Thus as such, these flowcharts are difficult to operationalize and study empirically (and in a predictive testing sense, impossible to test at their current stage of development).

Because of the problems involved in investigating these models, most initial "tests" of the models have been relatively simplistic, albeit relatively sophisticated statistically. Using examinations of the Howard-Sheth model as an example (Farley & Ring, 1970; Lehmann, Farley, & Howard, 1971; and Lehmann, O'Brien, Farley & Howard, 1971), several interesting observations are possible:

1. The "tests" have been largely cross-sectional.
2. The mathematical form used has been linear.
3. Parameter estimates have been made across people using regression (either OLS or TSLS).
4. The results are encouraging but mixed. The links in the cognitive side of the model, including such variables as brand comprehension, attitude, intention, and purchase have been both significant and plausible. The links among the informational variables, such as attention, perceptual bias, and overt search, on the other hand, have been much weaker.

At this point, many areas of the Howard-Sheth model are largely unexplored, including:

1. Non-linear forms
2. Lagged forms

3. Alternative operational definitions
4. Individual parameter estimates.

Thus the major characteristic of these general buyer behavior models is their limited operationalization.

Perceptual Mapping Models

Perceptual mapping models differ substantially from full-scale buyer behavior models. They focus on explanation of individual preference, and are not immediately concerned with either information reception on the one hand or choice on the other. The essential feature of these models is that they view brands as a collection of positions on a set of attributes, and preference toward a brand as some weighted combination of the positions.

Perceptual mapping models can be expressed graphically as in Figure 2. The essential postulate of these models is that the "closer" an alternative is to the ideal, the more preferred it is. In Figure 2, this would imply, assuming attributes 1 and 2 were equally important, that alternative A is the most preferred.

A variety of trends in the literature has suggested the perceptual mapping approach. In economics, Lancaster (1966) has proposed a utility theory based on the characteristics of a good instead of the good as a whole. The multidimensional scaling literature suggests that preference is a function of the distance of an object from the ideal (Green & Carmone, 1969; Kruskal, 1964 a & b; and Shepard, 1962 a & b). In social psychology, two very similar theories of attitude have been provided (Fishbein, 1967; and Rosenberg, 1956) which suggest that attitude is a weighted sum of positions on dimensions. Thus the essential concept that an object can be viewed as a point in n dimensional space is widely supported.

One reason why these research traditions were not merged sooner is the differences in terminology used to describe them. In order to make the similarity more obvious, the following glossary is useful:

<u>Construct</u>	<u>Alternative Designations</u>
The object to be rated	Object, Alternative, Brand
The dimensions on which the objects are related	Dimension, Characteristic, Attribute Goal
The position on the dimension	Position, Belief, Perceived Instrumentality
The weight of the dimension in the decision	Weight, Importance, Value, Saliency

With one important exception (Einhorn & Gonedes, 1971), all the perceptual mapping models proposed have been of the following form:

$$Y_j = \sum_{i=1}^n W_i |P_{ji} - I_i|^K$$

where W_i = weight of the i^{th} dimension

P_{ji} = position of the j^{th} object on the i^{th} dimension

I_i = ideal position on the i^{th} dimension

K = an integer

and n = number of relevant dimensions.

In other words, the attitude is a weighted sum of distance to the ideal on each of the relevant dimensions.

Depending on the way I_i and K are defined, Y_j can take on many forms. For example, a $K = 1$ implies city block distance while a $K = 2$ implies Euclidean distance. In two past tests, city block distance has proven best predictively (Bass, Pessemier & Lehmann, 1971; and Lehmann, 1971). However, other considerations, such as stability under orthogonal rotation (city block distance is not, while Euclidean is) and utility theory implications (city block implicitly assumes constant marginal utility on the attributes, while Euclidean is one form of diminishing marginal utility) may dominate in the selection of a distance measure. In any event, several alternative distance measures are available.

The relationship of distance to the ideal, similarity (which is the inverse of distance) to the ideal, attitude, and preference are also somewhat confusing. The relationship can be summarized as follows:

	<u>"Good" Object</u>	<u>"Bad" Object</u>
Distance	Small	Large
Similarity	High	Low
Attitude	High (favorable)	Low (unfavorable)
Preference	High (low in a ranking)	Low (high in a ranking)

Actually, preference is usually a comparative measure between attitudes, but for purposes of these perceptual mapping models, the two terms are largely synonymous. Thus the differences between the traditions in the literature are largely semantic.

A more fundamental reason why these traditions in the literature were not synthesized sooner is that there are essential differences between the models in terms of the way the dimensions arise. Two basic approaches exist, and they differ in some important features:

<u>Approach</u>	<u>Source of Positions on Dimensions</u>	<u>Position Scale</u>	<u>Distance Measure</u>
Multidimensional Scaling	Derived from Similarity Judgments	Metric	Minkowski (Usually Euclidean)
Direct Rating	Ratings on Pre-Specified Dimensions	Binary (Fishbein) or Metric	Minkowski (usually City Block)

Yet in spite of these differences, the essential similarity of these approaches is obvious.

Tests of these perceptual mapping models have differed substantially from those of the Howard-Sheth model. Because the mathematical form of the relation is pre-specified, the model is more operational. Efforts have centered around measurement of the three basic constructs (preference, weight, and position) and deducing the relevant dimensions.

In general, the tests have been greatly encouraging. Indirectly derived dimensions have proved useful in explaining preference among such diverse alternatives as automobiles (Green & Carmone, 1969), jobs (Hill & Pessemier, 1971), and political candidates (Johnson, 1970). Ratings on pre-specified dimensions have proved successful in analyzing such alternatives as television shows (Lehmann, 1971), soft drinks (Bass, Pessemier & Lehmann, 1971), and numerous branded products (Bass & Talarzyk, 1972; Ginter, 1972; and Winter, 1972). In all these examples, predictions based on a perceptual mapping model have greatly outperformed both demographics and random models.

In spite of these encouraging results, there are some important problems involved with applying perceptual mapping models. As stated, these models are largely tautological and as such can be investigated but not truly tested. Also attempts to use a subject-estimated ideal point have been disappointing (Bass et al; 1971; and Lehmann, 1971). Finally, the weights have not proved to be very useful (Lehmann, 1971; and Sheth & Talarzyk, 1972) for a variety of reasons (Beckwith & Lehmann, 1972). Thus substantial testing and refinement of these models is also needed.

A Synthesis

Looking at the pictorial representations of the two models of individual behavior (Figures 1 and 2), one is struck more by the differences than by the similarities. Considering the problems involved in testing either separately, the obvious question which arises is: "Why attempt to synthesize them?" The answer is that by combining them, both may benefit.

The obvious weakness of the perceptual mapping approach is that it does not suggest either how information influences the individual or how preference is related to choice. Placing it in the context of a general buyer behavior model suggests both. On the other hand, the obvious weakness of the buyer behavior approach is the imprecise formulation of the links between blocks in the flowchart. Using perceptual mapping makes some of the links both explicit mathematically and empirically viable.

To see how a perceptual mapping model might be combined, consider the Howard-Sheth model (Howard & Sheth, 1971), which is currently under revision (Howard & Ostlund, 1973), portrayed in Figure 1, and the perceptual mapping model represented by Figure 2. The perceptual mapping model can be viewed as a combination of four constructs: 1) Choice criteria, 2) Weights, 3) Brand comprehension, and 4) Confidence. The choice criteria can be viewed as the dimensions of the perceptual map and the weights as the weights attached to the dimensions. Brand comprehension could be treated as the position of the brand on the dimensions. Finally, the random component representing uncertainty suggested for introduction into the perceptual mapping models (Lehmann, 1971 a; & 1972 a & b) can be considered as a measure of confidence. In other words, the center of the Howard-Sheth model could be viewed as a perceptual mapping model.

Figure 1
HOWARD-SHETH MODEL

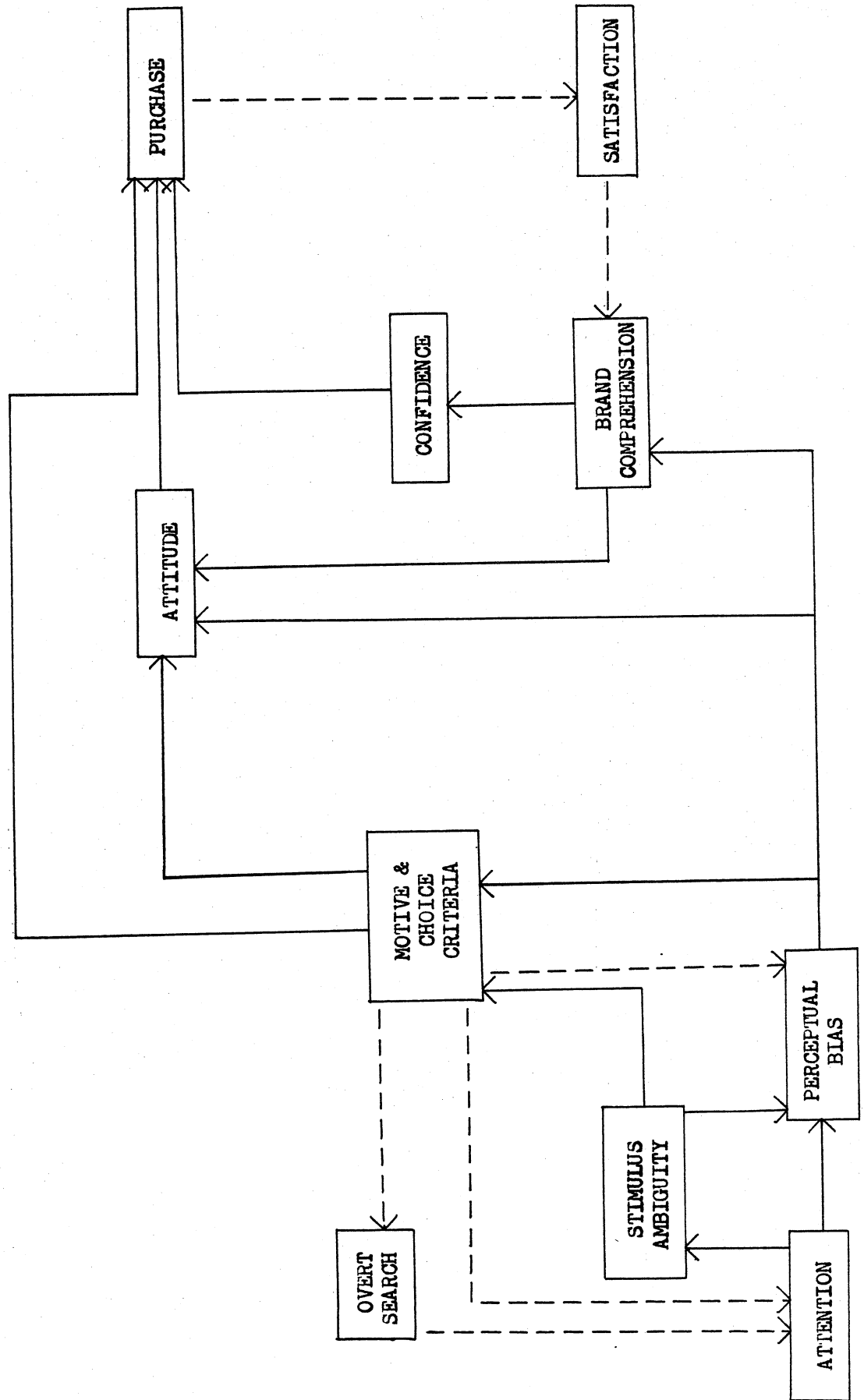
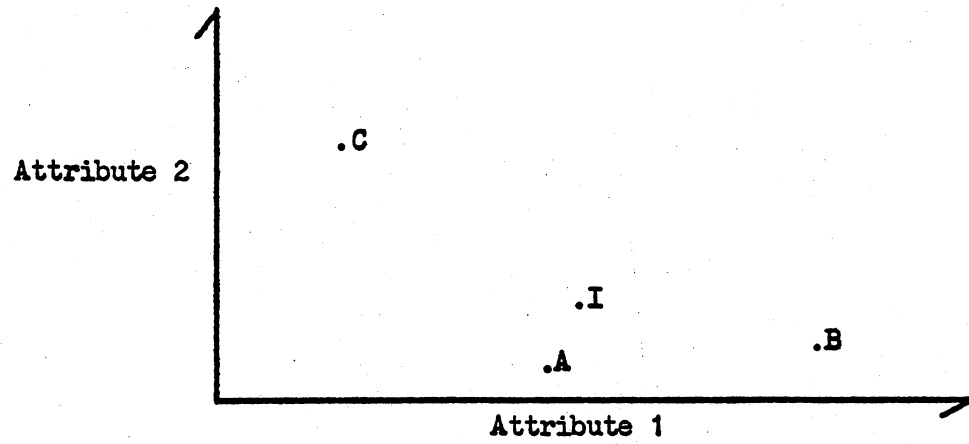


Figure 2

TWO-DIMENSIONAL EXAMPLE

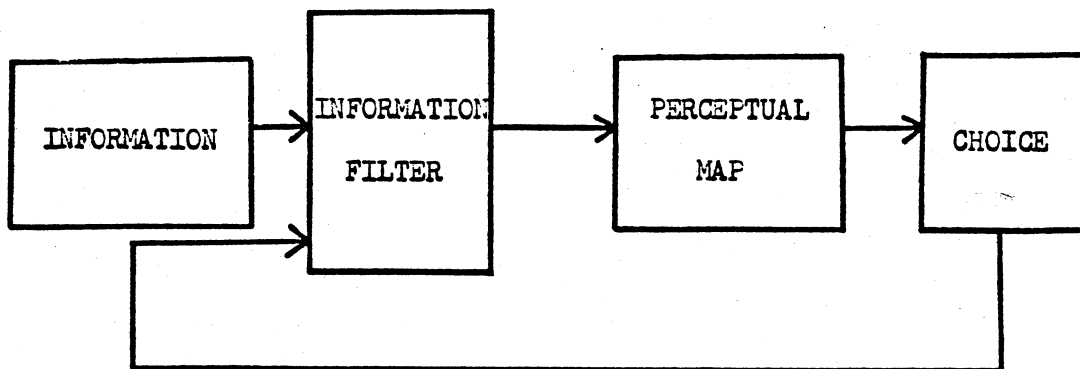


One advantage of this synthesis is the improved explanatory power for the center of the Howard-Sheth model. Another is that the "new" model can be expressed very simply (Figure 3). Possibly the most important advantage of the synthesis, however, is that it suggests how advertising and other information affects choice. Simply stated, this new model suggests that advertising influences choice by changing either the positions on the dimensions, the weights of the dimensions, or the ideal position on the dimensions. As such, it provides a structure for future research on the informational side of the model which has to date proved the most difficult to investigate.

Conclusion

Two major research traditions dealing with the individual consumer have emerged: Buyer Behavior Models and Perceptual Mapping. In spite of the problems involved in attempting to test (or at least investigate) these models, both have shown great promise. This paper suggests that combining the two traditions will be to the mutual benefit of both.

Figure 3
A SYNTHESIZED MODEL



References

- Bass, Frank M., Edgar A. Pessemier, & Donald R. Lehmann. An Experimental Study of Relationships Between Attitudes, Brand Preference, and Choice. Paper No. 307, Institute for Research in the Behavioral, Economic, and Management Sciences, Herman C. Krannert Graduate School of Industrial Administration, Purdue University, April, 1971.
- _____ & W. Wayne Talarzyk. An Attitude Model for the Study of Brand Preference. Journal of Marketing Research, 1972, 9, 93-96.
- Beckwith, Neil E. & Donald R. Lehmann. The Importance of Importances in Attribute Models of Consumer Preference. Working paper, Columbia University Graduate School of Business, 1972.
- Einhorn, Hillel J. & Nicholas J. Gonedes. An Exponential Discrepancy Model for Attitude Evaluation. Behavioral Science, 1971, 16, 152-157.
- Farley, John U. & L. Winston Ring. An Empirical Test of the Howard-Sheth Model of Buyer Behavior, Journal of Marketing Research, 1970, 7, 427-438.
- 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.
- Ginter, James L. An Experimental Study of Attitude Change and Choice of New Brands. Proceedings. Spring Conference, American Marketing Association, 1972.
- Green, Paul E. & Frank J. Carmone. Multidimensional Scaling: An Introduction and Comparison of Nonmetric Unfolding Techniques. Journal of Marketing Research, 1969, 4, 330-341.
- Hill, Raymond E. & Edgar A. Pessemier. Multidimensional and Unidimensional Metric Scaling of Preference for Job Descriptions. Paper No. 308, Institute for Research in the Behavioral, Economic, and Management Sciences, Herman C. Krannert Graduate School of Industrial Administration, Purdue University, April, 1971.
- Howard, John A. & Lyman E. Ostlund. Buyer Behavior: Theoretical and Empirical Foundations. New York: Alfred A. Knopf, Inc., 1973 (in press).
- _____ & Jagdish N. Sheth. The Theory of Buyer Behavior. New York: John Wiley & Sons, 1969.
- Johnson, Richard M. Political Segmentation. Marketing Review, 1970, 25, 20-24.
- Kruskal, J. B. Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis. Psychometrika, 1964a, 29, 1-27.
- _____. Nonmetric Multidimensional Scaling: A Numerical Method. Psychometrika, 1964b, 29, 115-129.

- Lancaster, Kelvin J. A New Approach to Consumer Theory. Journal of Political Economy, 1966, 74, 132-157.
- Lehmann, Donald R. A Study of Television Show Preference and Attributes of Programs. Proceedings. Spring Conference, American Marketing Association, 1972a.
- _____. Preference Among Similar Alternatives. Working paper, Columbia University, 1972b.
- _____. Evaluating Marketing Strategy in a Multiple Brand Market. Journal of Business Administration, 1971a, 3, 15-26.
- _____. Television Show Preference: Application of a Choice Model. Journal of Marketing Research, 1971b, 8, 47-55.
- _____. John U. Farley & John A. Howard. Testing of Buyer Behavior Models. Proceedings. Second Annual Conference, Association for Consumer Research, September 1-3, 1971c, 232-242.
- _____. Terrence V. O'Brien, John U. Farley, & John A. Howard. Empirical Contributions to Buyer Behavior Theory. Working paper, Columbia University, 1971.
- Nicosia, Francesco. Consumer Decision Processes: Marketing and Advertising Implications. Englewood Cliffs, N.J.: Prentice-Hall, 1966.
- Rosenberg, M. J. Cognitive Structure and Attitudinal Affect. Journal of Abnormal and Social Psychology, 1956, 53, 367-372.
- Shepard, Roger N. The Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function I. Psychometrika, 1962a, 27, 125-139.
- _____. The Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function II. Psychometrika, 1962, 27, 219-246.
- Sheth, Jagdish N. & W. Wayne Talarzyk. Perceived Instrumentality and Value Importance as Determinants of Attitudes. Journal of Marketing Research, 1972, 9, 6-9.
- Winter, Frederick W. Laboratory Experimental Study of Attitude Change and Brand Choice. Proceedings. Spring Conference, American Marketing Association, 1972.

COMPARISON OF METHODS FOR ANALYZING SOURCES
OF PERCEIVED DIFFERENCES IN PRODUCTS

Alan B. Flaschner
University of Georgia¹
and Lyndon E. Dawson, Jr.
Northeast Louisiana University

An attitude is the organization of concepts, beliefs, habits, and motives associated with a particular object or phenomenon (McKeachie and Doyle, 1966, p. 560). Consumer attitudes toward a product or service (an object or phenomenon) are viewed on the basis of several attributes which usually vary in terms of their relative "importance".²

In order to develop appropriate marketing strategies, a marketing manager may wish to know the relative importance of the factors that contribute to the overall perceived differences between his product and those of competitors. An understanding of the relative importance of the evaluative criteria used by consumers, then, is the core of a consumer-oriented approach to marketing (Engel, Kollat and Blackwell, 1968, p. 439). For maximum effectiveness, the marketing manager must design and market products in a manner consistent with the evaluative criteria consumers use (Engel, Kollat and Blackwell, 1968, p. 439).

The purpose of this research is to compare and evaluate alternative techniques which may be useful in judging the importance of product attributes.

Typical methods for obtaining measures of attribute importance include a dichotomous scale (Important-Not Important) for each attribute, rank ordering of the attributes, gradient scales, (e.g., 1-6) for each attribute, and point assignments from a common sum for each attribute (Schendel, Wilkie, McCann, 1971, p. 404).

The work reported in this paper is based on the assumption that the variation between stimuli, evaluated on a particular attribute, indicates the relative importance of that attribute. For example, in the tradition of multi-dimensional scaling it is assumed that the more important a criterion, the greater will be the respondent's ability to discriminate between stimuli on that criterion (Green and Carmone, 1970, p. 63).

Method

The problem, then, is to measure the "importance" of various attributes of a product without asking the direct question, "How important is that attribute?" The technique chosen to accomplish the task is magnitude estimation. Magnitude estimation is a procedure of free or unconstrained number matching. As was done in the present experiment the subject is given a reference point, or standard (a five percent alcoholic solution) which is assigned an arbitrary value such as 100. The subject indicates for each stimulus how it compares to the standard. For example, if, on a particular characteristic, the subject gives the stimulus a rating of 50, he is saying that it is half the standard on that characteristic; if, on a particular characteristic, he gives the stimulus a rating of 300 he is saying that it is 3 times the standard on that characteristic.

The assumptions underlying magnitude estimation are (1) that the range of numbers that can be assigned to a stimulus is unconstrained and (2) that estimation error will be random. That is, even though there may be wide variation between subjects, a subject's sense of the magnitude of a particular stimulus is valid. Given a random sample, it is assumed that if one subject overestimates another will underestimate; thereby, making the mean of the distribution the best estimate of the value of the stimulus.

Stimuli

The stimuli toward which the attitudes were measured were five (5) solutions of beer containing 5 percent, 7 percent, 9 percent, 11 percent, and 13 percent alcohol. To make sure that a relatively complete and unambiguous set of product characteristics were presented to the respondent (Schendel, Wilkie and McCann, 1971, p. 404) only attributes which fall within the cognitive-potency dimension (Engle, Kollat and Blackwell, 1968, p. 166) were used. "Strength, body, lightness, and sweetness" appear to be four mutually exclusive and all-inclusive "facts or beliefs" about beer.

Subjects and Procedure

The nineteen volunteer subjects, whose attitudes were measured, varied between the ages of 18 and 26. The subjects who received the solution in random order were asked to compare the solution to the standard on one randomly selected attribute. The subjects rated all five solutions (concentrations) on the randomly selected attribute before they rated the solutions on another randomly selected attribute.

Results

Table 1 is a summation of the subjects' perceptions of the five (5) concentrations of beer on four (4) attributes.

As indicated in Table 2 parts A, B, C, and D, using .05 level of significance as the criterion, the subjects were able to perceive differences among the five concentrations on the basis of strength and body, but were not able to do so on the basis of lightness and sweetness. Table 2 parts A, B, C, and D, using the .05 level of significance as the criterion, indicates that the variation among subjects was significant when they estimated the strength, body and lightness of the concentrations but was not significant when they estimated the sweetness of the concentrations. While it is acknowledged that the authors could have increased the sample size, without an increase in the variation among subjects, the increase in sample size would cause the variation among subjects to be significant.

Using the .05 level of significance as the criterion, Table 3, in which the experiment is treated as a complete factorial design (4 attributes x 5 concentrations x 19 subjects) with no replications and the three-way interaction term (attributes x concentrations x subjects) is used as the error term, indicates that there was a significant difference among attributes.

Determining Attribute Importance

The authors have shown that there was a significant difference among attributes. The relative importance of the attributes can be determined in several ways.

TABLE 1

PERCEPTION OF MAGNITUDE (AVERAGE ACROSS SUBJECTS, VARIATION AMONG SUBJECTS)

Attribute	% Alcohol	Geometric Mean	Standard Deviation For Geometric Mean*
Strength	5%	70.7	44.9
	7	105.3	68.6
	9	212.3	85.5
	11	273.2	133.1
	13	410.7	155.8
Body	5	92.1	59.2
	7	72.4	38.1
	9	122.2	69.5
	11	93.1	81.5
	13	121.6	102.3
Lightness	5	108.1	52.6
	7	106.0	47.7
	9	76.9	79.3
	11	49.9	103.2
	13	74.0	92.7
Sweetness	5	98.7	55.5
	7	68.5	44.3
	9	50.8	52.4
	11	29.9	35.4
	13	13.9	17.7

$$\text{*Standard Deviation For Geometric Mean} = \frac{\text{Standard Deviation For Arithmetic Mean} \times \text{Geometric Mean}}{\text{Arithmetic Mean}}$$

Kuepnapas and Wikstroem (1963) p. 618

TABLE 2
RANDOMIZED BLOCK DESIGN*

A: PERCEPTION OF STRENGTH						
<u>SOURCE OF VARIATION</u>	<u>DEGREES OF FREEDOM</u>	<u>SUMS OF SQUARES</u>	<u>MEAN SQUARES</u>	<u>F</u>	<u>SIG</u>	
Between Groups	22	2193088.0	99685.82	11.958	.814 x 10 ⁸	
Main Effects:						
Between Subjects	18	645564.8	35864.71	4.302	.272 x 10 ⁴	
Between Concentrations	4	1547523.7	386880.87	46.410	.284 x 10 ⁸	
Interaction (Residual)						
Between Subjects and Concentrations	72	600197.0	8336.07			
Total	94	2793285.0				
B: PERCEPTION OF BODY						
<u>SOURCE OF VARIATION</u>	<u>DEGREES OF FREEDOM</u>	<u>SUMS OF SQUARES</u>	<u>MEAN SQUARES</u>	<u>F</u>	<u>SIG</u>	
Between Groups	22	894504.0	40659.27	4.030	.249 x 10 ⁴	
Main Effects:						
Between Subjects	18	627412.1	34856.22	3.455	.193 x 10 ³	
Between Concentrations	4	267092.0	66773.00	6.619	.282 x 10 ³	
Interaction (Residual)						
Between Subjects and Concentrations	72	726369.0	10088.46			
Total	94	1620873				

TABLE 2 (Continued)

C: PERCEPTION OF LIGHTNESS						
SOURCE OF VARIATION	DEGREES OF FREEDOM	SUMS OF SQUARES	MEAN SQUARES	F	SIG	
Between Groups	22	978251.3	44465.97	3.826	.401 x 10 ⁴	
Main Effects						
Between Subjects	18	963454.1	53525.23	4.606	.150 x 10 ⁴	
Between Concentrations	4	14797.3	3699.33	.318	.865	
Interaction (Residual)						
Between Subjects and Concentrations	72	836767.7	11621.77			
Total	94	1815019.0				

D: PERCEPTION OF SWEETNESS						
SOURCE OF VARIATION	DEGREES OF FREEDOM	SUMS OF SQUARES	MEAN SQUARES	F	SIG	
Between Groups	22	127425.6	5792.1	1.23	.248	
Main Effects						
Between Subjects	18	87722.1	4873.4	1.04	.431	
Between Concentrations	4	39703.6	9925.9	2.124	.087	
Interaction (Residual)						
Between Subjects and Concentrations	72	337944.1	4693.7			
Total	94	465369.7				

* 19 SUBJECTS x 5 CONCENTRATIONS

TABLE 3

INCOMPLETE FACTORIAL DESIGN (SUBJECTS X CONCENTRATIONS X ATTRIBUTES) *

Source	DF	Sum of Squares	Mean Square	F-Ratio	Level of Significance
Between Groups	379	8021283.0	21164.34	2.605	.217 x 10 ⁷
Main Effects:					
Between Subjects	18	898148.6	49897.14	6.144	.103 x 10 ⁶
Between Attributes	3	1326714.2	442238.06	54.451	.483 x 10 ⁹
Between Concentrations	4	675560.7	168890.13	20.795	.390 x 10 ⁷
Two Way Interaction:					
Between Subjects and Attributes	54	1426047.8	26408.29	3.251	.427 x 10 ⁶
Between Subjects and Concentrations	72	746959.3	10374.43	1.277	.092 (N.S.)
Between Attributes and Concentrations	12	1193585.5	99465.44	12.247	.289 x 10 ⁶
Three Way Interaction (Residual):					
Between Subjects, Concentrations and Attributes	216	1754270.0	8121.61719		
Total	379	8021283.0			

*Of the 380 (19 x 5 x 4) Possible Data Points 369 Were Non Zero

One method of determining attribute importance is to examine the range of magnitude estimates. Discrimination theory leads one to believe that the perception of the difference between stimuli is a function of the relevance of that perception to the perceiver. Magnitude estimates, being ratio data, suggest that if the perceived difference between stimuli on one attribute is greater than the perceived difference between stimuli on another attribute and if this greater perceived difference between differences is caused by a greater importance of that attribute to the perceiver; then, by reverse logic, the attribute on which the concentrations are perceived to be most different is the most important to the respondent and the other attributes are correspondingly less important.

Welch (1971, pp. 76-87) suggests that the arithmetic mean consistently overestimates the central tendency of pooled ratio information. Using the geometrically averaged magnitude estimates to correct for this bias, Table 1 indicates that on the basis of: strength, the highest estimate is 5.82 times the lowest estimate; on the basis of body, the highest estimate is 1.68 times the lowest estimate; on the basis of lightness, the highest estimate is 2.16 times the lowest estimate; and on the basis of sweetness, the highest estimate is 7.1 times the lowest estimate.

This method of estimating attribute importance allows one to state the degree to which one attribute is more important than another. For example, based on Table 1, one can say that sweetness in beer is 1.22 (7.1/5.82) times as important to the respondents as is strength, that sweetness is 4.22 (7.1/1.68) times as important to the respondents as is body, that strength is 3.46 (5.82/1.68) times as important to the respondents as is body, etc.

The problem with this method of estimating attribute importance is that it fails to consider the expected monotonic relationship between the alcohol concentrations and the corresponding magnitude estimates.

A second method of determining attribute importance is to examine the degree to which the magnitude estimates for the various concentrations of alcohol fit the equation:

$$R = kS^n$$

"where R is the subjective, or response, magnitude in arbitrary units; S is the stimulus magnitude in physical units (alcohol concentrations of 5 percent, 7 percent, 9 percent, 11 percent and 13 percent); n is an empirical but non-arbitrary exponent; and k is an empirical constant dependent on the value assigned to the standard (100)" (Welch, 1971, p. 76). Because this simple power function has been found to hold for more than three dozen sensory continua (Stevens and Galanter, 1957), Welch (1971), Stevens (1957) and Shinn (1969) propose that the observed psychophysical relationship in which equal stimulus ratios produce equal perceptual ratios has achieved the status of a natural law (Stevens and Galanter, 1957, pp. 377-411; Welch, 1971, pp. 76-87; Stevens, 1957, pp. 153-181; Shinn, 1969).

By reverse logic, the attribute on which the concentrations are perceived to be most in accord with the power function is the most important to the respondents and the other attributes are correspondingly less important.

Table 4, in which is presented the amount of variation in the geometrically averaged magnitude estimates that can be explained by knowledge of the alcohol

concentrations, shows that the most important attribute is "strength," the next most important attribute is "sweetness" etc.

By using the degree to which the data fits the power function as a test of attribute relevance, we are able to show which attribute is more important but are unable to show the degree to which it is more important.

A third method for unobtrusively estimating attribute importance consists of running estimates of the "overall" perceived differences between the stimuli through a multistage nometric multidimensional scaling program. Since the first dimension that is derived from a multidimensional scaling program explains the majority of the variance between the products and the second dimension explains a little less of the variance, etc., the order in which the dimensions are listed indicates their degree of contribution to the "overall" perceived differences between the stimuli. A high degree of correlation between these calculated dimensions and the original attributes should indicate possible labels for the dimensions.

The influence of which solution was chosen as the standard was removed by using a normalization procedure based on the median (Thorndike, 1922, pp. 116-121) to rescale the data on each attribute.

Based on the assumption that a stimulus compared to a standard must bear a ratio relationship to another stimulus compared to the same standard in the present experiment the magnitude scalings were converted to paired comparison information (data that is appropriate for input to a multidimensional scaling program). The "overall" perceived difference between stimuli for an individual was calculated by using an expansion of the Pythagorean theorem for a 4 dimensional space. For an individual, the overall perceived difference between the two stimuli D_{jk} (say the .05 percent and .07 percent alcohol concentrations) was the square root of the squares of the perceived difference between the stimuli summed over the four attributes.

$$D_{jk} = \sqrt{\sum_{i=1}^n (X_{ij} - X_{ik})^2}$$

where: D = overall perceived difference between concentrations, given:

j = alcohol concentration
 k = alcohol concentration other than j

X = magnitude estimate, given:

i = particular attribute on which the magnitude estimate was made
 j = alcohol concentration
 k = alcohol concentration other than j

$n = 4$, the number of attributes on which the magnitude estimates were made.

The geometric average of the 19 D_{jk} , the "overall" estimates of the perceived difference between the stimuli, were used as input to TORSCA, a nometric multidimensional scaling program (Young, 1968, pp. 319-321). For the one dimensional solution (see Table 5), satisfactory stress of 0.0 was achieved. This means that one dimension explains the overall perceived difference between the concentrations and it makes little sense to discuss the two, three or four dimensional solutions.

To fit the power function, $R = kS^n$ all R values must be positive and greater

than 0.0. Retaining the distance between stimuli (retaining the interval between them), the configuration was revised by adding 1.0 to each value. As is indicated in Table 6, when the revised configuration is fit by the power function $R = kS^n$, 98 percent of the variation in R (the revised configuration) can be explained by knowledge of S (the alcohol concentrations).

Correlation between the TORSCA created values and the Thorndike (1922) re-scaled perceptions, indicates that the configuration is "strength" (see Table 7). The coefficient of stress in this configuration (0.0) indicates that strength and only strength was the attribute the respondents were considering when asked to evaluate concentrations on the basis of strength, body, lightness, and sweetness.

Conclusion

In an experiment in which the respondents estimated the magnitude of five alcohol concentrations of beer on four attributes the authors attempted to unobtrusively estimate attribute importance.

Based on the assumption that the perception of greater differences reflects greater importance, the ratio data revealed that the attributes could be ranked in terms of importance as follows: (1) sweetness, (2) strength, (3) lightness, (4) body. This method considered the degree to which one attribute was more relevant than another but ignored the expected monotonic nature of the data.

Based on the assumption that the better the magnitude estimates on a particular attribute fit the power function $R = kS^n$ the more relevant is that attribute, the ratio data revealed that the attributes could be ranked in terms of their importance as follows: (1) strength, (2) sweetness, (3) lightness, and (4) body. This method considered the monotonic nature of the data but ignored the degree to which one attribute was more relevant than another.

Rescaling the data, then collapsing the attributes into an "overall" estimate of the perceived difference between stimuli, multidimensional scaling revealed that the only attribute that was important to the subjects was "strength," the other attributes being unimportant. Because it considers both the degree of importance of the attributes as well as the monotonic nature of the data, multidimensional scaling seems to hold the most promise for unobtrusive attribute importance estimation.

Footnotes

1. Alan B. Flaschner is an Assistant Professor of Marketing at the University of Georgia and Lyndon E. Dawson, Jr. is a Professor of Marketing and Management and Coordinator of Graduate Studies in Business Administration at Northeast Louisiana University.
2. "Importance" is the desire or need for the presence of a particular attribute in a product (Schendel, Wilkie, and McCann, 1971, p.415).

References

- Engel, J. F., Kollat, D. T. & Blackwell, R. D. Consumer Behavior. New York: Holt, Rinehart and Winston, Inc., 1968.
- Green, P. E. & Carmone, F. J. Multidimensional Scaling and Related Techniques in Marketing Analysis. Boston: Allyn and Bacon, Inc., 1970.

- Kuennapas, T. & Wikstroem, I. Measurement of Occupational Preferences: A Comparison of Scaling Methods. Perceptual and Motor Skills, 1963, 17, 611-624.
- McKeachie, W. J. & Doyle, C. L. Psychology. Reading, Mass.: Addison Wesley Publishing Co., Inc., 1966.
- Schendel, D. E., Wilkie, W. L. & McCann, J. M. An Experimental Investigation of "Attribute Importance". Proceedings of the 2nd Annual Conference, Association for Consumer Research, 1971.
- Shinn, A. M., Jr. The Application of Psychophysical Scaling Techniques to Measurement of Political Variables, Working papers in Methodology No. 3. Chapel Hill: University of North Carolina, Institute for Research in Social Science, 1969.
- Stevens, S. S. On the Psychophysical Law. Psychological Review, 1957, 64, 153-181.
- Stevens, S. S. & Galanter, E. Ratio Scales and Category Scales for a Dozen Perceptual Continua. Journal of Experimental Psychology, 1957, 54, 377-411.
- Thorndike, E. L. An Introduction to the Theory of Mental and Social Measurement. New York: Teachers' College, Columbia, 1922.
- Welch, R. E., Jr. The Use of Magnitude Estimation in Attitude Scaling: Constructing a Measure of Political Dissatisfaction. Social Science Quarterly, 1971, 52, 76-87.
- Young, F. W. TORSCA, an IBM Program for Nometric Multidimensional Scaling. Journal of Marketing Research, 1968, 5, 319-321.

AN INDIVIDUAL DIFFERENCES MULTIDIMENSIONAL SCALING APPROACH
TOWARDS CONSUMER DECISION-MAKING¹

Richard I. Kushner
University of New Hampshire

The study described herein was conceived in the context of a particular theoretical viewpoint about consumer behavior which is based on two major assumptions. First, it is assumed that a consumer perceives a class of products in terms of multiple attributes. Thus, product perceptions may be represented by a multidimensional space. Second, it is assumed that differences in consumer's preferences and consequent purchases are based on differences in assessment with respect to the various perceptual dimensions or on differences in weightings of the dimensions as bases for forming preferences. The primary objective is to examine this particular approach to the study of consumer preferences and decision-making.

The study had three major procedural steps which operationalize how the prime objective was to be achieved. First, Ss were asked to rate cars with respect to various specified characteristics with the expectation that the major characteristics of perceptions of cars would be thereby manifested, directly or indirectly. These characteristics-ratings were examined for evidence bearing on the assumption of differences in perceptual style. Specifically, Ss in this study were asked only to rate how much of a characteristic they felt a car possessed. They were not asked to scale on the basis of how important any characteristic was to them, i.e., what Ss actual point of view was toward a characteristic. The usage of perceptual style then refers to scaling of the first type, i.e., how much of a characteristic a car has. Second, the Ss were also asked to rate the cars with respect to preference, and these preference ratings were examined for evidence of preference styles. Third, the characteristics-ratings were examined in the light of the preference ratings to find out whether preferences and characteristics-ratings were systematically related.

In perceptual research the term dimensions frequently refers to factors which are mutually orthogonal or linearly independent. In the present study no such analysis of the characteristics was attempted so that to be consistent dimensions will be referred to only in the theoretical sense. If the characteristics-ratings had been transformed into mutually orthogonal factors, then the factors which emerged could legitimately be referred to as dimensions. The point is that the characteristics used reflect concerns similar to those in using "pure" dimensions.

The goal of research in this area could involve the elaboration and systematization of descriptions of human behavior in decision-making by referring to obtained differences along a specified number of relevant dimensions. It should be made clear that the present research is only exploratory and thus should be seen as a first step in the direction of the theoretical viewpoint mentioned above.

It is felt that consumer behavior models could account for significantly more variability in the behavior of interest by incorporating the multidimensionality of the stimuli. Any product can be defined along a number of attributes, and the decision made by any consumer should reflect or maintain some functional relationship to the set of attributes used. The full perceptual

space for a given class of products consists of all dimensions with respect to which those products could be discriminated by consumers, but for a given consumer, the functional product stimulus can be completely defined in terms of what dimensions are attended to, i.e., what dimensions are given nonzero weights.

Even if two consumers attend to the same dimensions, their perceptions of a given product may differ because of differences in their evaluations of that product with respect to some of the attended dimensions. Furthermore, even if two consumers attend to the same dimensions and agree in their evaluations of a given product with respect to those dimensions, they may differ in the strengths of their preferences for that product because they attach different weights of importance to the attended dimensions. (Note that this involves two types of scaling procedures and that for this research, SS were asked only to rate how much of a characteristic a car contained and not how important that characteristic was.) The following is an example of multidimensional perception. If one wants to purchase a car, one may consider durability, prestige, dependability and style. Here are then four characteristics which could define my perception of a given car. One may view car x as a highly durable and dependable car, whereas another individual may view the same car as stylish and prestigious. The same stimulus is construed by the two people on different characteristics and thus their perception of this car will be different. Two other individuals may both consider car x an extremely prestigious car, but for the first individual prestige has a high positive weighting and for the second individual prestige has a high negative weighting. Their respective decisions about preference will be different based on knowledge of just this single characteristic. The individual who thinks car x is prestigious and considers a good car to be prestigious (positive weighting) makes one choice decision. The second individual who thinks the car is prestigious but dislikes prestige in a car (negative weighting) will make a different choice decision. Thus their decisions may be a function of the characteristics they are concerned with in terms of their relative weightings of importance in the decision-making. If one wants to account for differences in decisions made by consumers, one must investigate the dimensions that are potentially relevant for a stimulus product and determine how important those dimensions are for a given specified sample of consumers. This will involve individual multidimensional scaling of the two sorts described above.

A model is needed which takes into account any existing relationship between a consumer's preference for a product and his perception of that product. Individuals whose pattern of preferences are alike can then be described on the basis of their ratings of cars on a specified number of characteristics. Individuals whose preference patterns are different will also rate cars differently on the characteristics. An interaction is being predicted between preference and characteristics-ratings. It is the pattern of characteristics-ratings which emerges for a group of individuals who agree in their preferences which defines a perceptual style. The question becomes, are one's preferences a function of his perception operationalized in terms of a multiplicity of rated characteristics defining any given stimulus product? Are one's preferences reflected in the pattern of ratings of products on these characteristics?

The multidimensional approach has been used in situations different from the ones discussed here, i.e., consumer decision-making. Its successful use in other areas of decision-making should serve to show its utility as a viable paradigm. The present concern is to account for variability in one's preferences by determining what characteristics correlate strongly with a consumer's preferences. In the case of cars, for example, if a particular con-

sumer rated a set of cars extremely high on preference, and at the same time considered those cars to be extremely durable, this relationship might indicate that durability is a positively weighted dimension which is crucial in determining a given consumer's preference for car name stimuli. (He might also rate nonpreferred cars low in durability.) Wiggins and Wiggins (1968) were concerned with accounting for men's preferences for females by utilizing three dimensions, those being size of breasts, buttocks, and legs. Their approach involved having Ss make paired-comparison preference ratings of nude female silhouettes which varied along the three dimensions. In making decisions about feminine attractiveness the authors concluded that these three dimensions were crucial factors in determining preferences. Thus depending upon how important one considers each of the three dimensions, a researcher might be able to predict an individual's preferences for females. The kinds of questions asked by Wiggins and Wiggins and the model that they utilized are similar to those of this research project.

The basic concern of this research is to determine the relationship between Ss relative preferences for each of a set of cars and ratings of the cars on a given set of characteristics. If a given set of Ss rate a number of cars for preference and then rate those same cars on a number of characteristics used to describe cars, will consistent trends or relationships emerge? Ss have been grouped according to their preferences for cars and their ratings of a set of cars on characteristics. The relationship between preferences for the stimuli and the ratings on characteristics have also been examined.

The position taken is that consumers are concerned with a number of dimensions of any stimulus product and that some are weighted more heavily in terms of importance than others in accounting for consumer decision-making. One car is preferred over another because (1) of a set of underlying dimensions of concern, and (2) of where the person places himself on those underlying dimensions in terms of their importance to him in his decision-making. The present research will obtain scaling data concerned with the first point only. A more complete definition of perception and/or perceptual style would involve scaling to examine both components. With this research one has only a first approximation in that direction.

Method

Subjects

The Ss were 40 male and 40 female students enrolled in introductory psychology classes at the University of New Hampshire; course credit was given for participation as Ss.

Apparatus and Procedure

The names of 20 cars served as stimuli. The 20 cars were judged the most familiar out of a set of 53 in an independent pilot study. Car names were used as the stimuli because cars were presumed to have high interest value to the sample of Ss being used. On the first meeting of the present study, Ss were asked to rate the 20 car names on a 0 to 100 percent preference scale. This scale was selected because of its familiarity in terms of the numbers and percentiles. A 0 percent rating indicated a car that S did not like or admire, would not desire to own, and thus would not choose out of a list of cars. A 100 percent rating meant that the car was one which S considered to possess to a maximum all those qualities which would make a car absolutely preferred. An 'x' mark on the scale indicated a response as to where Ss felt a particular

car name belonged according to his or her own preference. An 'x' mark could go anywhere on the scale, either between the numbers or directly over a number. Ss were instructed to rate the car names according to interval scale properties.

One week later the Ss were asked to rate each of the 20 car names on each of 12 characteristics. A 0 to 100 percent scale was used for this purpose. Ss, for example, were asked to judge how prestigious car x was and to judge how economical or stylish car y was. For any given characteristic a rating of 0 percent meant that for that S, in his opinion and according to his standards, a particular car did not in any way possess the quality of that characteristic. 0 percent thus meant that a given car for a given characteristic was not at all representative of that characteristic. A 100 percent rating meant that a particular car was maximally representative of the characteristic being considered.

The 12 characteristics used were dependability, durability, size, economy, power, horse power, style, quality of ride, handling of car, prestige, luxury and resale value. These 12 were selected in a pilot study in which Ss were asked to rate a list of characteristics in terms of their importance in making preference ratings about cars. Ss were also instructed to write down any other characteristics they felt were important that were not on the list. Based on these results, 12 characteristics were selected. Each of the 12 characteristics were explicitly defined for the Ss so that they could refer to them throughout the experiment. (See appendix.)

Reliability and transitivity of the preference ratings by comparable Ss were checked in pilot work. The preference ratings were reliable and with few exceptions, transitive. Only four of 40 pilot Ss showed violation of transitivity. Ss for this pilot work were grouped in terms of low, medium and high familiarity with the 20 car names in general. For one randomly selected S from each group, the correlation of preference ratings made one week apart was computed. The three correlations were .84, .76, and .86.

Results

Comparison of Ss with respect to preferences utilized transpose principal components analysis of preference ratings followed by varimax rotation. The concern was to determine if there were consistent differences on the basis of preferences. Each S factor indicated a pattern of preferences across the 20 cars that differed from those of other S factors. Results showed that 68 percent of the Ss could be classified in one of three groups. Ss were classified on the basis of which factor they loaded most highly on. For these data 15 factors were produced based on the criterion of having eigenvalues greater than one. The first three factors together accounted for 53 percent of the total variance. Table 1 presents for each preference factor the mean factor loadings for each subgroup of Ss.

Table 2 presents the rank orderings of the 20 cars according to the mean preference ratings of subgroups of each of the three preference groups. Each of the first two subgroups consists of the 10 Ss who loaded most highly on the corresponding S factor. The third subgroup consisted of 5 Ss who loaded most highly on the corresponding S factor. The third subgroup consisted of 5 Ss who loaded most highly on the third preference factor. The concern was to obtain groupings of Ss as alike in their preferences as possible. An examination of this table shows the differences in preference ratings for each of three homogeneous subgroupings of Ss who serve as exemplars for the three preference factors. It can be noted that for Preference Group 1 small sports type cars, e.g., Volkswagon Beetle, Toyota Corona, and Saab, are ranked very low in preference, whereas for Preference Group 2, these same cars are ranked

Table 1

Mean Loadings of Homogeneous Subgroups of
Ss for each Preference Factor

	F ₁	F ₂	F ₃
10F ₁ <u>Ss</u>	.84	.09	-.04
10F ₂ <u>Ss</u>	.08	.82	.02
5F ₃ <u>Ss</u>	-.07	.05	.70

near the top. In Preference Groups 1 and 2 the Rolls Royce, Buick Riviera, and Lincoln Continental are in the upper 25 percent on preference whereas in Preference Group 3 they fall within the lower 30 percent. Also of interest from Table 2 is that while Volkswagon Beetle and Saab are ranked at the bottom (18th and 20th) for Preference Group 1, and near the top (5th and 7th) for the Preference Group 2, they are ranked at about the center of the distribution (10th and 11th) for the Preference Group 3 Ss.

A transpose principal components analysis was applied to Ss' ratings on the 12 characteristics. Each of 82 percent of the Ss could be classified by highest loading in one of three groups. For these data 10 factors emerged, and the first, second, and fourth factors together accounted for 60 percent of the total variance. The third factor was omitted since only a very few Ss loaded high on it.

A chi square test of independence showed that the three preference-pattern groups and the three characteristics-pattern groups were significantly related ($\chi^2 = 33.29$, 9 df, $p < .001$; See Table 3). A second chi square test showed that sex and characteristics-ratings pattern were also significantly related ($\chi^2 = 7.55$, 2 df, $p < .05$; See Table 4).

A multiple regression solution was obtained for the three Preference Groups. The criterion was preference averaged across members of the Preference Group, and the predictors were mean ratings on the 12 characteristics; the 20 cars constituted the sample. Using a multiple regression scheme is one way of characterizing the preference groupings. It provides a descriptive way of differentiating preference homogeneous subgroups. An examination of Tables 5, 6, and 7 presents the correlations of the characteristics with each other and with preference for each of the three homogeneous preference groupings. It should be noted that the characteristics are quite highly correlated for all three Preference Groups which makes interpretation difficult.

The multiple correlation for Preference Group 1 on preference was .99. While 12 predictors were used, it can be shown that many of the predictors could be dropped from the system and still have a multiple correlation that is near .99. For example, arbitrarily take the highest correlation between preference and one of the characteristics as presented in Table 5. The power characteristic, for example, correlates .96 with preferences. Using the semi-partial approach to multiple correlation ($R = r^2_1 + r^2_{(2 \cdot 1)}$), we could next obtain the correlation for a second characteristic (e.g., durability) with

Table 2

Ranking of Preference Groups on the Basis of Means
Across Homogeneous Subgroups of S_s for Each Car

Preference Group One	\bar{X}	Preference Group Two	\bar{X}
1. Rolls Royce	87.50	1. Mercedes Benz	90.00
2. Mercedes Benz	83.00	2. Rolls Royce	86.00
3. Chevrolet Corvette	81.00	3. Triumph	81.00
4. Lincoln Continental	79.00	4. Lincoln Continental	70.50
5. Pontiac Grand Prix	66.00	5. Saab	67.00
6. Buick Riviera	63.00	6. Chevrolet Corvette	67.00
7. Mercury Cougar	62.00	7. Volkswagon Beetle	66.50
8. Ford Country Squire	58.00	8. Volkswagon Carman Gia	65.50
9. Pontiac Catalina	55.50	9. Buick Riviera	48.50
10. Buick Skylark	54.50	10. Toyota Corona	46.00
11. Chevrolet Malibu	51.00	11. Pontiac Grand Prix	46.00
12. Ford Galaxie	50.50	12. Mercury Cougar	39.50
13. Triumph	48.00	13. Pontiac Catalina	38.50
14. Ford Mustang	42.00	14. Ford Mustang	38.00
15. Chevrolet Nova	41.50	15. Buick Skylark	34.00
16. Plymouth Valiant	29.00	16. Chevrolet Malibu	29.50
17. Toyota Corona	25.50	17. Ford Galaxie	29.00
18. Saab	17.00	18. Ford Country Squire	28.50
19. Volkswagon Carman Gia	14.50	19. Plymouth Valiant	23.00
20. Volkswagon Beetle	14.50	20. Chevrolet Nova	22.50

Preference Group Three	\bar{X}
1. Mercedes Benz	73.00
2. Chevrolet Nova	61.00
3. Toyota Corona	61.00
4. Chevrolet Malibu	59.00
5. Ford Country Squire	54.00
6. Triumph	49.00
7. Buick Skylark	48.00
8. Ford Mustang	46.00
9. Chevrolet Corvette	46.00
10. Saab	45.00
11. Volkswagon Beetle	42.00
12. Mercury Cougar	36.00
13. Ford Galaxie	34.00
14. Rolls Royce	32.00
15. Plymouth Valiant	30.00
16. Pontiac Grand Prix	28.00
17. Volkswagon Carman Gia	27.00
18. Pontiac Catalina	22.00
19. Buick Riviera	19.00
20. Lincoln Continental	17.00

Table 3

Chi Square Test for Independence Between
Preference and Characteristics Patterns

		OBSERVED				
		Preferences				
		F ₁	F ₂	F ₃	F _{other}	
Characteristics	F ₁	2	14	5	5	26
	F ₂	9	1	0	7	17
	F ₃	8	2	4	4	18
	F _{other}	1	3	1	8	13
		20	20	10	24	74

Table 4

Chi Square Test for Sex

		OBSERVED		
		Male	Female	
Characteristics	F ₁	18	8	26
	F ₂	5	12	17
	F ₃	35	26	18
				61

preference with power partialled out of durability. These two resulting correlations can be squared and summed which yields the proportion of variation in preferences for the first grouping accounted for on the basis of just two predictors, i.e., power and durability. The results of these manipulations yielded a multiple correlation of .98. It can be seen that just two predictors yield an extremely high multiple correlation accounting for most of the variability in the criterion. While in this example the semi-partial does not add a large portion of accountable variability to the already large power-preference correlation, the example is used to illustrate the approach used with this model.

The multiple correlation for Preference Group 2 on preference was .98. Resale value correlates .94 with preference and a semi-partial approach could be used to search for any other single characteristic which could add significant accountable variability. A second predictor would be needed which correlated highly with preference and only slightly with resale value. A single predictor yielding $r = .94$ makes negligible the need to find a specific additional predictor which could boost the multiple correlation close to .98.

Table 5

Correlations of Mean Car Ratings for Preference Group One

	Ch 1	Ch 2	Ch 3	Ch 4	Ch 5	Ch 6	Ch 7	Ch 8	Ch 9	Ch 10	Ch 11	Ch 12
Pref	0.88	-0.87	0.96	0.93	0.94	0.92	0.96	0.94	0.94	0.84	0.82	0.66
Ch 1	--	-0.67	0.90	0.76	0.79	0.89	0.90	0.91	0.89	0.85	0.94	0.55
Ch 2	--	--	-0.88	-0.93	-0.88	-0.73	-0.85	-0.84	-0.84	-0.65	-0.57	-0.66
Ch 3	--	--	--	0.95	0.95	0.89	0.98	0.97	0.96	0.86	0.83	0.64
Ch 4	--	--	--	--	0.97	0.82	0.93	0.90	0.91	0.79	0.71	0.63
Ch 5	--	--	--	--	--	0.83	0.96	0.91	0.91	0.86	0.77	0.56
Ch 6	--	--	--	--	--	--	0.89	0.90	0.91	0.80	0.84	0.66
Ch 7	--	--	--	--	--	--	--	0.96	0.96	0.91	0.86	0.58
Ch 8	--	--	--	--	--	--	--	--	0.99	0.83	0.84	0.70
Ch 9	--	--	--	--	--	--	--	--	--	0.82	0.83	0.70
Ch 10	--	--	--	--	--	--	--	--	--	--	0.89	0.24
Ch 11	--	--	--	--	--	--	--	--	--	--	--	0.37
Ch 12	--	--	--	--	--	--	--	--	--	--	--	--

Ch 1--Dependable

Ch 5--Style

Ch 9--Quality of ride

Ch 2--Economical

Ch 6--Durable

Ch 10--Handling of car

Ch 3--Powerful

Ch 7--Prestige

Ch 11--Resale value

Ch 4--Horsepower

Ch 8--Luxury

Ch 12--Size

Table 6
Correlations of Mean Car Ratings for Preference Group Two

	Pref	Ch 1	Ch 2	Ch 3	Ch 4	Ch 5	Ch 6	Ch 7	Ch 8	Ch 9	Ch 10	Ch 11	Ch 12
Pref	--	0.83	0.15	0.53	0.33	0.60	0.84	0.57	0.45	0.57	0.85	0.94	-0.08
Ch 1	--	--	0.31	0.32	0.08	0.32	0.96	0.34	0.32	0.46	0.62	0.88	0.00
Ch 2	--	--	--	-0.67	-0.77	-0.55	0.30	-0.66	-0.70	-0.58	0.21	0.15	-0.76
Ch 3	--	--	--	--	0.93	0.90	0.35	0.98	0.94	0.93	0.43	0.52	0.59
Ch 4	--	--	--	--	--	0.87	0.12	0.91	0.86	0.81	0.35	0.31	0.55
Ch 5	--	--	--	--	--	--	0.37	0.92	0.78	0.76	0.60	0.57	0.33
Ch 6	--	--	--	--	--	--	--	0.38	0.34	0.47	0.66	0.91	0.00
Ch 7	--	--	--	--	--	--	--	--	0.93	0.92	0.48	0.56	0.56
Ch 8	--	--	--	--	--	--	--	--	--	0.95	0.24	0.43	0.77
Ch 9	--	--	--	--	--	--	--	--	--	--	0.37	0.55	0.66
Ch 10	--	--	--	--	--	--	--	--	--	--	--	0.82	-0.36
Ch 11	--	--	--	--	--	--	--	--	--	--	--	--	-0.05
Ch 12	--	--	--	--	--	--	--	--	--	--	--	--	--
Ch 1---Dependable						Ch 5---Style					Ch 9---Quality of ride		
Ch 2---Economical						Ch 6---Durable					Ch 10---Handling of car		
Ch 3---Powerful						Ch 7---Prestige					Ch 11---Resale value		
Ch 4---Horsepower						Ch 8---Luxury					Ch 12---Size		

Table 7

Correlations of Mean Car Ratings for Preference Group Three

	Pref	Ch 1	Ch 2	Ch 3	Ch 4	Ch 5	Ch 6	Ch 7	Ch 8	Ch 9	Ch 10	Ch 11	Ch 12
Pref	--	-0.05	0.39	-0.27	-0.44	-0.01	0.02	-0.29	-0.29	-0.36	-0.34	-0.22	-0.32
Ch 1	--	--	-0.47	0.69	0.53	0.24	0.88	0.74	0.69	0.65	0.70	0.83	0.42
Ch 2	--	--	--	-0.81	-0.92	-0.71	-0.26	-0.81	-0.89	-0.91	-0.81	-0.62	-0.93
Ch 3	--	--	--	--	0.90	0.67	0.47	0.96	0.93	0.88	0.92	0.80	0.74
Ch 4	--	--	--	--	--	0.72	0.31	0.90	0.90	0.89	0.90	0.67	0.84
Ch 5	--	--	--	--	--	--	-0.01	0.66	0.64	0.61	0.55	0.22	0.66
Ch 6	--	--	--	--	--	--	--	0.49	0.46	0.45	0.52	0.73	0.18
Ch 7	--	--	--	--	--	--	--	--	0.94	0.88	0.92	0.79	0.74
Ch 8	--	--	--	--	--	--	--	--	--	0.96	0.93	0.78	0.87
Ch 9	--	--	--	--	--	--	--	--	--	--	0.91	0.76	0.92
Ch 10	--	--	--	--	--	--	--	--	--	--	--	0.85	0.77
Ch 11	--	--	--	--	--	--	--	--	--	--	--	--	0.55
Ch 12	--	--	--	--	--	--	--	--	--	--	--	--	--

Ch 1--Dependable	Ch 5--Style	Ch 9--Quality of ride
Ch 2--Economical	Ch 6--Durable	Ch 10--Handling of car
Ch 3--Powerful	Ch 7--Prestige	Ch 11--Resale value
Ch 4--Horsepower	Ch 8--Luxury	Ch 12--Size

For Preference Group 3 from factor three on preferences there is no approximation of a multiple regression system, the reason being that upon examination of Table 7 it becomes apparent that the correlations between characteristics are larger than the correlations of preference and characteristics. The problem is that with such a pattern of correlations one is likely to obtain an inflated multiple correlation which is spurious. This is known as the suppressor effect and results when predictors correlate slightly with the criterion but highly with each other. Table 7 clearly indicates that the correlation of preference with a characteristic is lower than that characteristic with most other predictors.

One can still examine differences in the three Preference Groups on the basis of examining the correlations between preference and characteristics must be guarded because as stated previously, the 12 characteristics used are not orthogonal but rather in most cases correlate quite highly with each other. The reader is referred to Tables 5, 6, and 7 for the following discussion.

For Preference Group 1 Ss, economy correlates strongly but negatively with preference (-.87), for Group 2 Ss, it seems not to be a relevant characteristic in determining their preferences (.15), while for Group 3 Ss, economy correlates positively with preference (.39). Here is a characteristic in reference to which the groups can be differentiated and characterized. For Preference Group 1 Ss a car that has economy is low on the preference hierarchy, for Group 2 Ss it is a characteristic which correlates very slightly with preference, indicating it is not searched for by these Ss when rating cars for preference, and for Preference Group 3 Ss a car possessing economy is high on the preference hierarchy.

Economy and horsepower are the two characteristics which correlate most strongly with preference for Preference Group 3 Ss, economy correlating positively and horsepower negatively. For Preference Group 1 Ss the signs of the correlations are reversed, with economy being negatively correlated with preference and horsepower positively. On this basis Preference Group 1 Ss could be characterized as preferring those cars which they consider to possess a great deal of horsepower and not much economy, while for Preference Group 3 Ss a highly preferred car is one which they deem to be high on economy and low on horsepower. It is interesting to note that for both Preference Group 1 and Preference Group 3 Ss horsepower and economy are negatively correlated, -.94 and -.92, respectively.

Power and horsepower correlate strongly and positively for Preference Group 1 Ss while they correlate only moderately for Preference Group 2 Ss, accounting for over three times as much variance in Preference Group 1 Ss. For Preference Group 3 Ss they correlate negatively. Referring back to Table 2, it can be seen that cars rated high in preference for Preference Group 1 Ss contain quite a bit of power and horsepower while cars rated high in preference for Preference Group 3 Ss contain little horsepower and power. Note that for Preference Group 3 Ss cars rated low in preference contain quite a bit of power and horsepower (e.g., Lincoln Continental, Rolls Royce, etc.)

The size of a car appears to be a characteristic which helps to differentiate the preference groupings. A car's size correlates positively with preference for Preference Group 1 Ss, practically not at all for Preference Group 2 Ss, and negatively for Preference Group 3 Ss. The size of a car matters for Preference Group 1 and 3 Ss in determining their preferences (but in opposite ways).

Quality of ride correlates strongly and positively with preference for Preference Group 1 Ss and negatively for Preference Group 3 Ss. Of interest is that while resale value is a highly desired characteristic for Preference Group 1 and 2 Ss, it correlates negatively with preference for Preference Group 3 Ss.

Discussion

It has been shown that Ss can be differentiated on the basis of their preference ratings and also on the basis of characteristic ratings. It can also be concluded that there is a systematic relationship between perceptual style and preference patterns. It was expected that, since several separate factors emerged on the basis of the preference ratings, unique predictive or descriptive systems would be necessary to characterize homogeneous groups of Ss for each factor. Results confirmed these expectations since a single prediction system could not account for the variability in the preferences as well as distinct prediction systems can. Since there are different Ss' factors on the basis of preference, different predictive schemes will be necessary in order to characterize different subgroups of people.

No multiple regression statement could be made for Preference Group 3. It could be that their preference patterns are not linearly or additively related to characteristics ratings. It could be that the relations between preferences and characteristics for this group are curvilinear. A curvilinear regression system might be needed which could further be broken down into linear, quadratic, cubic, etc. components. For example, it might be the case that Preference Group 3 Ss prefer cars with a medium amount of prestige, but do not prefer cars with a great deal or very little prestige. With a multiple linear regression system, correlations between preference and prestige would be low and not considered predictive when, in fact, with some other system it would. In any case future research is needed to gather more information to adequately describe this preference style.

It is clear from the results that the characteristics of products used are descriptive of the different preference styles which emerged from the principal components analysis. Variability in car preferences can be accounted for by a multidimensional approach even though in the present case the characteristics used were not orthogonal dimensions. In the present correlational setting, characteristics were strongly related to preferences and thus useful in describing the homogeneous Ss' groupings and also suggestive of this approach's potential utility in developing predictive systems sensitive to individual differences on the construct of concern. Future research is needed to determine what characteristics of products and people are most important in consumer decisions for any particular homogeneous grouping of Ss relative to the behavior of interest and to gather more information about Ss which could serve to differentiate them on that particular behavior. This added information could form the basis for new predictors in a descriptive system. In this sense the present research was only exploratory, concerned with testing and evaluating a particular model of decision-making which emphasizes individual differences in multidimensional perception.

Many individual difference dimensions of possible utility in a predictive system were not included because it was not known at the start of this project whether different factors of Ss would emerge on preference ratings. If only one factor emerged so that only one prediction scheme were needed for all Ss, it would have been fruitless to collect a great deal of data on individual difference variables since there would have been no differences in preference

to account for. A basic objective of this research was to determine whether there exist relatively distinct patterns of preference, each of which is characteristic of reasonably substantial numbers of Ss. This now being known, new research designs can be created taking this into account and can begin to develop and test predictive systems for different groups of like Ss.

The implication is that separate and distinctive descriptive systems will have to be developed to account for the individual differences that could develop on the dependent variable of interest. If one predictive system only is created in an attempt to account for all the variability in the behavior of interest, without breaking this variability up on the basis of like performing subgroups of Ss, interactions of Ss with treatment conditions could very possibly attenuate or eliminate the phenomenon under investigation. For example, if no attempt in the present research were made to differentiate Ss on the basis of preferences, correlations of the characteristics with preferences would almost certainly be low due to interactions across Ss. One set of Ss might rate a Rolls Royce high on prestige and preference while a second set of Ss might rate a Rolls Royce high on prestige and low on preference. It should be clear that by attempting to predict across both types of Ss, the correlations would be reduced when in fact if separate predictive systems were used on the basis of homogeneous groups of Ss, correlations might be higher and predictions more accurate.

Hopefully the research paradigm approximated in this research will be used by other researchers to collect data as to the validity of such an approach in asking and answering questions about consumer decision-making. The model appears to be broad and flexible in scope, capable of providing the researcher with a methodology to examine questions about human behavior. The model involves the gathering of data on Ss and then identifying Ss on the basis of similar profiles across the behavior of interest. Homogeneous subgroups of Ss could be formed on the basis of factor analytic procedures. The next step involves determining what the relevant person variables are that interact with treatment conditions. The concern is to identify those relevant person variables which allow one to describe, predict, and account for the variability existing in the dependent variables of concern. Do those Ss who fall into the same subgroupings have similarities on personality, perceptual, and cognitive individual difference dimensions? Autobiographical data can also be used which might also prove helpful in defining the homogeneous subgroupings. Ss across the different subgroupings could be compared on the basis of all the person measures available.

It has been stated that more information is needed to help further differentiate and define homogeneous subgroups with like preference profiles. As more information can be gathered on Ss, a beginning can be made to refine the different predictive systems necessary for different subgroups of people on preference. What types of information could be beneficial and how could it be used to create predictors within the present correlational methodology? It has been stated that one problem with this study is that Ss were only asked to scale on the basis of how much a car had of a given characteristic and not to scale how important that characteristic was in a decision. Could these two types of scaling procedures if utilized together, yield predictive information that would further help differentiate Ss on preference? Ss could be asked to not only rate a car for how much of a given characteristic it possesses but also as to how important that characteristic is for him in making a preference judgment. This type of information could serve a predictive value in accounting for more variability.

In the present multiple regression system the 12 characteristics served as predictors. It could be the case that a predictor had a high positive weight based on Ss rating cars as having a great deal of the particular characteristic when in fact Ss felt it to be a neutral one, i.e., not of concern when making preference judgments about cars. In this case the emerging positive weight is spurious. For example, Ss might consider a set of highly preferred cars to have a great deal of style and based on this single predictor, would be given a high positive weight. It might be the case, however, that these Ss feel that style is of no value or has a zero rating on a scale of importance from $\frac{-5 \quad 0 \quad +5}{\quad}$, where a -5 rating indicates a characteristic which matters in a negative sense and a +5 rating indicates high positive importance when regarding preference judgments. These two measures could be treated as predictors of preference in a multiple regression system.

There is an interaction between these two predictors in that a car rated as possessing a little style could be highly preferred (if style has a negative weight) or not preferred (if style has a positive weight). On the other hand a car rated as possessing a great deal of style could be highly preferred (if style is of positive importance) or not preferred (if style is of negative importance). To include these measures as two independent predictors in a multiple regression prediction system would be to lose variability that could be accounted for by their interaction. The implication is that a derived score is needed, taking into account both predictors, i.e., a single composite score. The assumption is that one can make better predictions if Ss can be differentiated on a new second measure. This new composite predictor would be different than the two alone in a multiple regression scheme, i.e., it would take into account the interaction and maximize the predictions of preferences. For example, an S might give style a -4 rating for importance and a 90 percent rating for how much style a given car has. The composite predictor would reflect the negative evaluation of importance by simply multiplying the two together.

With the above analysis in mind, consider how this second predictor could aid in defining preference groupings in the multiple regression data presented in this study. Based on ratings of how much a particular car possesses a given characteristic, one might obtain a homogeneous cluster of individuals all rating the characteristics in the same way. Upon consideration of the second predictor, this grouping can be broken down and analyzed still further into at least three subgroupings. They would involve, for example, rating style as desirable, as unimportant, and as undesirable, i.e., it does matter, but in a negative way. In all cases each subgroup rates the cars as having a lot of style.

If one of the goals of research is to truly describe homogeneous subgroups of individuals alike on the behavior of interest, then this type of derived score method might differentiate Ss more accurately. The implication is that while more information is needed to describe the preference groupings, just adding predictors, i.e., characteristics, will not be enough, but rather derived measures will be needed which will take into account interactions between the characteristics. Just adding more predictors is not the answer, but creating derived measures utilizing person dimensions of relevance and the two types of scaling procedures described could aid greatly in the development of valid predictive schemes.

Implications for Future Research

This research broadly defined has been dealing with an information-pro-

cessing approach to decision-making with a stress on the multidimensionality of the stimulus and the need for individual differences measures. The implications are broad sweeping. Far too often in psychology today the nature of the input has been denied with the result that many conclusions drawn are incomplete or do not adequately represent the complexity of the behavior being examined. It must be remembered, as Garner (1970) states, that there could be no processing of information, no decisions made, without an input to the system. Is it not time to turn interests towards defining the input and isolating variables relevant to its nature or composition? Future research needs to address itself as to the role of the stimulus in behavior and decision processes. As Garner states (1970), "for too long we have considered that a stimulus is a stimulus is a stimulus, whose only function is to elicit behavior. But all stimuli are not equivalent, and all information cannot be processed the same way (p. 357)." The present exploratory research is an attempt to define what the relevant dimensions are for a particular stimulus domain.

The present research has implications for advertising and marketing. An advertising organization could select a sample of consumers from a population of interest, i.e., potential consumers, and have them rate their product(s) for preference or desirability. This same sample could then rate the product(s) on a number of dimensions of relevance to that product. Prior research would be needed to determine what those dimensions are. Consumers in the sample could also rate how important each of the dimensions are to them for making a decision about a particular product. This data could then be analyzed and projections made to the population of interest. Those resulting perceptual styles which were related most systematically to the high preference ratings could be used as the basis of advertising those products. Those dimensions which accounted for most of the variability in preferences could be used as the basis for creating an image most viable for promoting that particular product. If, for example, advertisers found that durability were a highly weighted positive dimension, it could be built into the image. Those dimensions which were weighted high but negative could be handled in a parallel manner. Trial and error advertising could be reduced with a more profitable investment of funds in a sound research project.

Different styles of life might have different perceptual styles with corresponding different relationships to preferences for a given product. Thus the same advertising image for a product might not be fruitful for one group whereas it is for another. For example, a businessman planning to open up a car dealership in a rural area may want to know what type of dimensions are highly relevant there as opposed to an urban environment. It might be that economy is much more important in a rural area and prestige and power more important in an urban region. The dimensions he would want to key on would be different, and this is valuable information to possess. This same individual may decide to sell Volkswagon in a rural area but Corvettes and Cadillacs in an urban region based on such information. What this research has tried to demonstrate is that more precise and refined information about the product itself is of value theoretically and in applied industrial areas.

A limitation of this research has been the homogeneous sample of Ss. This restricted the range of responses on the dimensions and on preferences.

Footnotes

1. I wish to express my gratitude to Drs. Leslie A. Fox, G. Alfred Forsyth and Daniel C. Williams for their encouragement, direction and assistance in the preparation of this thesis.

References

- Cooley, W., & Lohnes, P. Multivariate procedures for the behavioral sciences. New York: Wiley, 1962.
- Coombs, C. H. Inconsistency of preferences: a test of unfolding theory. In W. Edwards and A. Tversky (Eds.), Decision-making. Maryland: Penguin Books, Inc., 1967, 319-332.
- Edwards, W. The theory of decision-making. In W. Edwards and A. Tversky (Eds.), Decision-making. Maryland: Penguin Books, Inc., 1967, 65-96.
- Garner, W. The stimulus in information processing. American Psychologist, 1970, 25, 350-358.
- Harman, H. H. Modern factor analysis. Chicago: University of Chicago Press, 1967.
- Horst, P. Factor analysis of data matrices. New York: Holt, 1971.
- Lee, W. Decision theory and human behavior. New York: Wiley, 1971.
- Neidell, L. The use of nonmetric multidimensional scaling in marketing analysis. Journal of Marketing, 1969, 33, 37-43.
- Siegel, S. Nonparametric statistics. New York: McGraw-Hill, 1956.
- Torgerson, W. S. Theory and methods of scaling. New York: Wiley, 1967.
- Tucker, L. R. Systematic differences between individuals in perceptual judgments. In G. Bryan and M. Shelly (Eds.), Human judgments and optimality. New York: Wiley, 1964.
- Wiggins, J., Wiggins, N., & Conger, J. Correlates of hetero-sexual somatic preference. Journal of Personality and Social Psychology, 1968, 10, 82-90.
- Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.
- Woodworth, R., & Schlosberg, H. Experimental psychology. New York: Holt, 1963 (1938).

Appendix

1. Dependable--reliable transportation without breaking down
2. Economical--low cost for fuel, low cost for upkeep
3. Powerful--a car that in your judgment represents social influence and authority
4. Horsepower--the purely physical power of the car
5. Style--the design of the car in terms of outward appearance
6. Durable--to stand up for rugged service despite hard use and abuse
7. Prestige--status assigned to an individual in a social group by owning a particular car
8. Luxury--car that possesses features which offer the most physical comfort and satisfaction although not necessary for maintaining the car
9. Quality of ride--smooth, comfortable ride under all possible weather conditions
10. Size--spaciousness of car
11. Resale value--the amount of money that can be received for a car when one wants to sell it
12. Handling of car--stability of car while taking curves and corners

THE FUTURE OF BUYER BEHAVIOR THEORY¹

Jagdish N. Sheth
University of Illinois

The motivational force behind writing this paper can only be the need for venturesomeness. This need was aroused partly by the speculative nature of the topic itself, which is to predict the future of buyer behavior theory, and partly by the deft encouragement of the session chairman, Jerry Zaltman. I am quite certain that some of my colleagues will disagree with things I foresee in buyer behavior theory, while others are likely to approve my assertions about the future of buyer behavior theory. Like a good forecaster let me simply say that "I may be mistaken but I am never wrong."

Future predictions typically entail the utilization of the Bayes theorem in some way because essentially prognosis implies revising the prior probability based on the assessment of some symptoms currently manifested in the phenomenon to be forecasted. Future predictions are also typically hazardous to one's welfare in any discipline because often the prior probabilities are no better than random probabilities due to relatively short histories of the disciplines, and because the assessment of contemporary symptoms is very difficult due to the rapid and complex changes the discipline may be undergoing. Buyer behavior theory may very well present these problems.

My objective in this paper accordingly is to estimate the prior probability by very briefly reviewing the historical perspective of buyer behavior discipline and then focus on a number of current events which are likely to determine both the velocity and the direction of buyer behavior theory. I shall focus on changes in structure and content of the theory which are likely to arise during the current decade and, so to speak, go out on a limb in my speculative tree.

A Brief Historical Perspective of Buyer Behavior

In the last quarter of a century, in my opinion, we have come a long way from the dark ages of strictly sporadic and random research in buyer behavior. The cumulative research effort in buyer behavior, both academic and professional, theoretical and empirical, or published and unpublished, is indeed impressive as can be gauged from several recent reviews (Guest, 1962; Howard, 1965; Twedt, 1965; Burk, 1967; Sheth, 1967; and Perloff, 1968). A closer examination of these reviews clearly indicates that we can identify four distinct phases of differential thoughts and emphases in the discipline.

The Empirical - Inductive Phase

The decades of the thirties and forties seemed to be dominated by

strictly empirical research mostly conducted by or for the industry's marketing decisions and their impact in the market place. Furthermore, the major emphasis tended toward gauging the effects of distribution, advertising and promotion decisions.

Among the several distinct characteristics of this phase, we may list (1) dominance of economic theory of the firm and especially the concepts of monopolistic competition, marginal utility analysis, and welfare economics; (2) macro market analysis at the aggregate level or at best at some pre-defined segmented level; and (3) emphasis on market's behavior responses as opposed to psychological responses.

The only exception to the above characterization of this phase of buyer behavior discipline seemed to be the acceptance of motivation research in which both the concepts and the methods of clinical psychology were widely applied to the understanding of buyer behavior.

The Formative Phase

The decade of the fifties must be regarded as the formative years of buyer behavior in which several major elements cemented the foundation of buyer behavior theory.

The first such element was the shift from measurement of aggregate to individual buyer behavior. Two different groups of scholars simultaneously contributed toward bringing about this change. The first was the Lazarsfeld School of Sociologists interested in measuring total change in voting behavior based on longitudinal panels which led to the establishment of household consumer panels in buyer behavior. The second was the Katona School of Economic Psychologists interested in building better indicators of economic growth based on the micro data of household acquisitions and inventories of durable appliances.

The availability of data on household purchase behavior eventually led to the interest in developing quantitative measures of brand or store loyalty and switching behaviors which in turn brought the utilization of stochastic processes such as Markov chains.

A second major element of the fifties was the growing interest in providing explanations for buyer behavior differences based on the social environment of the consumer. This led to the borrowing of the concepts of social stratification, reference groups, role orientations and opinion leadership. The major discipline relied upon, therefore, tended to be sociology and economic anthropology. Even though the theorizing was un-systematic and less refined in these attempts, they should be regarded as the pioneering efforts in search for causal explanations from disciplines other than the economic theory.

The third distinct element was the initial introduction to formal model building of buyer responses to marketing stimuli based on the optimization theory of operations research and econometrics (see Bass, et.al., 1961 for examples). Simultaneously, the utility theorists in

economics were also formalizing Samuelson's revealed preference theory of consumption. The combined effect was the introduction of deductive formal model building based on normative axioms and assumptions. This was further facilitated by the feasibility of building complex simulation models with the use of the computer. Given the infant stage of buyer behavior theory development, it is quite surprising to observe that management science was adopted in buyer behavior so early in the game. It is therefore, not at all surprising that most efforts at normative-deductive model building met with failures and premature rejection. As we shall see later, this fact has had tremendous implications for the future of buyer behavior theory.

It must be pointed out that all through these formative years, the empirical research on buyer behavior continued to accelerate independently due to the availability of micro data and the computer facilities to analyze them.

The Middle-Range-Theory Phase

The first half of the sixties can best be described as the identification stage. Buyer behavior began to be understood for its own sake rather than from the point of view of the marketer or the government or some other entity with vested interests. In my opinion, this can be directly attributed to the intensive borrowing of theories and concepts from those branches of behavioral sciences which had emerged as "pure" disciplines themselves. This included experimental psychology with emphasis on learning and perception, social psychology with emphasis on cognitive consistency, and rural sociology with emphasis on adoption processes of innovative products, practices and services. For further discussion I must refer you to the excellent review provided by Burk (1967).

The intensive borrowing from the behavioral sciences by numerous researchers, each one interested in some aspect of buyer behavior and predisposed or trained in some branch of behavioral sciences, resulted in the development of well-identified middle-range theories of buyer behavior. Any examples must include Howard's learning theory, Bauer's perceived risk theory, and several researchers developing theories based on Festinger's cognitive dissonance, Lewin's field theory, opinion leadership, and innovativeness and even on several personality theories. As I pointed out elsewhere (Sheth, 1967), the outstanding characteristic of this borrowing phase was the partial explanations each theory provided to the otherwise complex phenomenon of buyer behavior especially the one related to the problem-solving and habitual buying decisions. Not very surprisingly, other types of buyer behaviors were neglected in these middle range theories including unplanned impulsive behavior, novelty-seeking and situationally-anchored behaviors because very little theorizing was offered by the behavioral sciences in these areas.

Even though the bulk of this era concentrated on building middle-range theories, the efforts to build formal models based on optimization theory including linear programming, on stochastic processes including Bernoulli and Markovian processes, and on heuristics and other Monte Carlo type techniques continued. In fact, the early sixties can be regarded as

the golden era of management science in buyer behavior as evidenced from Massy, Montgomery and Morrison (1970).

The Integrative-Comprehensive-Theory Phase

The last half of the sixties and early years of the seventies is best identified with the emergence of comprehensive theories of buyer behavior. This basically entailed integrating several middle-range theories which had come to be accepted as well as putting together empirical research not identified with any theory in buyer behavior. (Nicosia, 1966; Howard and Sheth, 1969; Andreason, 1965; Engel, Blackwell & Kollat, 1968; Sheth, 1971; Sheth, 1972). In my opinion, the integrative-comprehensive theory building brought three factors in the development of the discipline. First, it emphasized the limitations of direct borrowing of theories from behavioral sciences without first adapting them to the complexity of buyer behavior. Furthermore, it established a precedent in reversing the process of borrowing by first conceptualizing the buyer behavior phenomenon and then searching for as many constructs as can be logically found in behavioral and social sciences. Second, and perhaps most important, the integrative-comprehensive theories brought to bear in buyer behavior the self confidence of independently building theories of buyer behavior in place of simply applying a social science theory to buyer behavior area with or without modifications. This must be regarded as the genesis for the emergence of buyer behavior as a discipline in itself rather than simply a problem area which can be explained by some social science discipline. Finally, they provided insights into building complex but realistic formal models of buyer behavior which may have contributed toward changing the traditional course of model building in terms of starting with simple, unrealistic assumptions and relaxing them to make them realistic as was true, for example, in utility theory (Katona, 1953).

During this phase, two other developments in buyer behavior theory are worth noting. The first was discarding the deductive-normative model building approach based on operations research methods in favor of statistical inductive model building with the use of multivariate analysis of large scale survey data. The second development was the broadening of marketing and buyer behavior horizons to nontraditional areas such as population control, nutrition, and public service delivery systems under the pioneering efforts of Kotler and Zaltman.

Assessment of History and Contemporary Signals

From the above brief historical review, it is my contention that buyer behavior theory is a far cry from random thinking. Within a very short period of time, we seem to have firmly laid the foundation for building a distinct discipline of buyer behavior which will neither be a subsystem of marketing nor that of any of the other older social sciences. Even more pleasant observation is that we seem to have achieved better

with respect to richness of thinking, comprehensiveness of theorizing, and testing of theories in naturalistic and realistic settings than many of the older behavioral science disciplines in their comparative periods of development. And why should it not be that way? Unless the newer disciplines learn to avoid the trial-and-error learning of older disciplines similar to the experiences of developing nations, there is very little hope of uplifting ourselves from becoming strict problem-solvers for the government or the industry. In summary, it is safe to predict that buyer behavior theory has nothing but a bright future.

However, an assessment of a number of contemporary events also indicate that the velocity of growth will not only be more rapid and diffused but is likely to significantly change the course of the growth curve. Rather than listing my assessment of these events and then forecasting the future directions of buyer behavior theory, I plan to devote the rest of the paper in detailing the major dimensions of future direction of the theory and link them to my assessment of contemporary events. The future developments are described below in terms of the following categories: (1) structural changes in buyer behavior theory, (2) broadening the horizons of relevance and applications of buyer behavior principles, and (3) active interest of other disciplines to borrow from buyer behavior theory.

Structural Changes in Buyer Behavior Theory

In the Bayesian crystal urn, I foresee four types of structural changes in the development of buyer behavior theory during the decade of the seventies. They are (1) establishing criteria to evaluate the relevance of different theories of buyer behavior, (2) constructing tests and scales to measure widely accepted hypothetical constructs in buyer behavior theory, (3) building complex formal but highly realistic and inductive models based on comprehensive theories of buyer behavior and, (4) theoretical and empirical research on nonpurposeful buyer behavior hopefully leading to a comprehensive theory.

Criteria for Evaluating Theories of Buyer Behavior

Although we have developed several theories of buyer behavior in both middle-range and comprehensive categories, surprisingly we have so far failed to develop widely accepted criteria with which to evaluate their usefulness or even relevance to buyer behavior. To be sure, existing theories are differentially accepted and diffused, but this seems to be based on tenuous factors such as the degree of face validity or predictive validity, the reputation of the author, the prestige of the institution, and ability to generate more research funds and comensurately more publicizing of the theory. In view of the fact that even greater number of researchers are likely to contribute in the coming years, I think the discipline is likely to experience personal rivalries and

showmanship among competing authors unless some evaluative criteria are developed and accepted. I foresee three different types of criteria emerging within the next five years in order to minimize the personal rivalry and showmanship mentioned above.

The first type of criteria seems already to be emerging in the form of the development of a meta theory of buyer behavior (Zaltman, Pinson, & Angelmar, 1972). In other words, theories of buyer behavior are likely to be critically examined, compared, and contrasted based on some fundamental judgments of philosophy of science. This type of criteria are strictly discipline-oriented and tend to emphasize the evaluation of the process of theorizing.

The second type of criteria, on the other hand, is likely to be pragmatic by being based on the usefulness of a theory of buyer behavior to solve specific problems. The emphasis in this type of evaluation is likely to be on the capability of a theory to enable the problem solver in achieving his own goals. What are the entities who are likely to increasingly utilize buyer behavior theories to help solve their problems? Obviously, the public policy makers, the business managers and consumer advocates seem the most likely entities. To the extent that each entity differs in its own perspective and activity, we are likely to see very different ways by which each entity is likely to put buyer behavior theories to its own use. For example, the research based on buyer behavior theories is likely to be very specific, ad hoc and symptomatic in the case of consumer advocates since typically they have tended to be issue-oriented. The public policy makers, on the other hand, are more likely to utilize comprehensive theories to conduct research on an exhaustive and systematic basis to search for the root causes of problematic symptoms pointed out by the consumer advocates. At least this is my hope. Finally, the marketing management is certainly likely to favor comprehensive theories of buyer behavior to understand and monitor market behavior simply because past experience has given enough evidence that middle range theories are not satisfactory.

The final type of criteria is likely to emerge from the efforts to generalize a theory to diverse and nontraditional areas of buyer behavior. The greater the ability of a theory to extend itself with a minimum number of modifications to the unexplored areas of buyer behavior such as search for information process or to nontraditional areas such as product utility and value formation, the greater is likely to be its popularity and diffusion. However, the diffusion of the theory is likely to be evaluated in terms of specific criteria similar to the criteria in statistics for parameter estimation procedures.

It is my hope that with the development of agreed-upon criteria in buyer behavior, we will see greater rigor and deductive logic in future theories of buyer behavior.

Standardized Measurement of Buyer Behavior

Even though we are still striving to develop better theories of buyer

behavior, I think we have a consensus on several hypothetical constructs in buyer behavior. These include the constructs of brand and store loyalty, behavioral intentions or buyer plans, predispositions toward choice alternatives, and perceptual biases in selective exposure and processing of information just to name a few. In addition, there seems to be a basic understanding that individual differences in buyer behavior are likely to be determined by constructs such as the life cycle, life style, socioeconomic status and role orientation differences among consumers. Secondly, we have recently tended to follow the psychometric tradition of data analysis especially with the use of multivariate methods which has brought to our attention the need for better and isomorphic measures of the above mentioned constructs which we strive to relate to one another in order to describe and explain buyer behavior.

I, therefore, foresee major research effort in buyer behavior channeled toward developing standardized scales for many of the constructs in buyer behavior mentioned above. I also think that this research effort is likely to be heavily influenced by the psychometric theories of scaling. Several important implications emerge from this forecast. First, we are likely to become more skeptical of the direct use of standardized sales and tests developed for comparable constructs in behavioral sciences. For example, in the area of attitudes, it is more likely that we will question the measurement procedures proposed in expectancy-value models in social psychology. Thus, we are likely to separate theories from measurement as we continue to borrow from the behavioral and social sciences. Second, the buyer behavior theory is very likely to become more mathematical and formal due to the research thrust in measurement and development of tests for the constructs. Finally, the development of standardized tests is likely to augment the empirical research in the unexplored areas of buyer behavior because research efforts will tend to be routinized as it is true today in some branches of psychology.

Quantitative Modeling of Buyer Behavior

In the distant future, I foresee reemergence of quantitative model building in buyer behavior. In other words, it will be quite some time before good mathematical models of buyer behavior are likely to emerge. Furthermore, the model building effort is likely to be distinctly different from what has been historically attempted in marketing. First, the models are likely to be problem-oriented instead of technique-oriented. Thus, by definition, they will be empirical in nature summarizing the efforts to research a problem area with the use of existing theories of buyer behavior. For example, a number of researchers are currently applying various theories of buyer behavior to understand how advertising works, how public delivery systems can be made more efficient in health, education and welfare, and how future transportation and communication needs can be fully met without endangering the environment. Second, the models are likely to utilize several statistical techniques in some sort of sequential multistage process rather than try to fit the empirical problem into a single technique such as mathematical programming or multidimensional scaling. Finally, the quantitative models of buyer behavior will be more positive rather than normative. In other words, we are more likely to see

predictive models of buyer behavior and less likely to see control models. Accordingly, the model building effort will coincide with testing and continuous updating. Furthermore, the adaptive control concepts are likely to be at the core of the updating process. In short, Bayesian philosophy is likely to dominate the model building effort replacing the search for optimality.

I also think that a number of researchers will attempt to decompose agreed-upon comprehensive theories of buyer behavior into smaller theories and develop models for them. We have already seen some efforts in this direction in regard to information processing and attitude structure subsystems of buyer behavior.

Research on Nonpurposeful Behavior

Based on the historical review, it is fair to state that we have so far concentrated on habitual, purposeful and problem-solving buyer behavior. Furthermore, we have developed several fairly comprehensive theories to explain it. However, the nonpurposeful behavior has received relatively little attention in the past, and it is, therefore, likely to emerge as the major substantive area of empirical research and theory building. By nonpurposeful behavior, I am primarily referring to curiosity, novelty seeking and exploratory behavior as well as cue-triggered impulsive buyer behavior. Some theorizing on this aspect of buyer behavior based on Berlyne's theory is recently developed by several researchers (Howard and Sheth 1969, Hansen 1972, Venkatesan 1972). However, considerable work still remains to be done especially in terms of gathering empirical data before a good systematic theory of nonpurposeful behavior can be developed.

It seems inevitable but to speculate that any comprehensive theory of nonpurposeful behavior will be extremely difficult to build and when built, it will more heavily depend on the typology of situational influences surrounding the nonpurposeful behavior than on the personal attributes of the buyer.

Broadening the Horizons of Buyer Behavior

Simultaneously with the structural changes in buyer behavior theory, I foresee rapid applications of buyer behavior concepts to three broadly-defined substantive problem areas. These are (1) cross-cultural buyer behavior research required by the emergence of multinational corporations, (2) public policy research on marketing institutions and practices required by increased concern in consumer welfare on the part of regulatory agencies, and (3) extension of buyer behavior to nontraditional areas of societal problems directly attributed to mass consumption nature of our society. I expect bulk of applied research in buyer behavior limited to these problem areas in the coming decade.

Cross-Cultural Buyer Behavior

I foresee extensive applied research in buyer behavior across different cultures simply because of the increased multinational characterization of most large business corporations. With the maturity of many foreign markets, the multinational corporations have become sensitive to the marketing-orientation in their efforts to diffuse products and services on a worldwide basis. The need to understand the impact of cultures on buyer behavior is obvious. In fact, it is surprising to observe how little attention has been paid to assess cultural influences on buyer behavior despite the fact that the United States is recognized to be the melting pot of diverse ethnical groups.

I think research on cross-cultural buyer behavior is likely to go through the same stages as what buyer behavior theory itself has gone through. First, there will be clusterings of studies mostly localized to some industries and some countries. Given the pattern of multinational business expansion, it is logical to presume that cross-cultural studies will be concentrated in European markets and with respect to nondurable consumer goods. Second, we should expect the development of several middle-range theories of cross-cultural buyer behavior based on the research in the first stage. Furthermore, the middle-range theories will be generated by the marketing practitioners and not by the academicians due to substantial costs involved in cross-cultural research. Of course, the academic scholars are likely to be instrumental in guiding the marketing practice's efforts to build middle-range theories. Finally, I foresee the eventual emergence of several comprehensive theories of buyer behavior. These will mostly constitute efforts to integrate both diverse middle-range theories and existing empirical research in cross-cultural buyer behavior. Furthermore, I also think that there will be virtually no differences between cross-cultural theories and domestic theories of buyer behavior in their structure.

Public Policy Research

It is but inevitable to foresee rapid applications of buyer behavior theories for the purpose of better regulation of marketing practices and institutions. Perhaps the single most factor for this observation is the singular inadequacy of the traditional concepts of micro economic theories to properly guide the function of regulation. A related reason is the divorce of marketing from economics during the past two decades. A second major factor is the increased pragmatism recently exhibited by the regulatory agencies such as the FTC and the FDA. This has brought home the need for empirical research on buyer behavior at the micro level both before and after major regulatory decisions, for example, the policy of corrective advertising, in order to ensure that desired consequences follow from them. Finally, the recent emergence of consumerism is likely to encourage the public policy makers to conduct fundamental research on buyer behavior in the hopes of producing good legislative policies.

The bulk of research in buyer behavior for public policy is likely to be problem-oriented, and most of the problems are likely to emerge from the negative side effects of mass marketing and mass consumption, for example, mass media effects on the citizen's values.

Buyer Behavior Research on Social and Environmental Problems

Perhaps the most critical applications of buyer behavior theories are likely to emerge from the research on social and environmental problems. We have already witnessed some utilization of marketing and buyer behavior concepts to social problems of less developed economies, for example, population explosion and malnutrition. However, the societal and environmental problems directly related to mass consumption and mass production are closer to home for most people working in consumer behavior. These problem areas include environmental and social pollution, welfare of minorities, and delivery of public services such as education and health. The greater social consciousness in solving these problems fortunately seems to be transcending the vested interests of the components of our mass production and mass consumption system so that research in these areas will tend to be nonpartisan.

In addition to the broadening of the horizons of buyer behavior, I foresee two subtle and indirect benefits arising from the research on social and environmental problems. The first is the separation of buyer behavior theory from marketing theory. In other words, I foresee emergence of greater respectability of buyer behavior theory by its extension to socially relevant issues. It is surprising, on reflection, to see how much psychology has suffered from the crisis of relevance because it emerged as a discipline in those areas of research which were not considered relevant or essential to mass consumption societies.

The second benefit is the rapid cross-fertilization of philosophy, theory and methodology between natural sciences and buyer behavior theory. I think it is simply inevitable that we will be working together with researchers from hard sciences such as physics, mechanics and biochemistry in search for solution of social and environmental problems. Thus, rather than borrowing from the other social sciences, it is likely that we will be borrowing the philosophy and methodology from the hard sciences. To me, this appears to be a unique opportunity for buyer behavior theory to elevate itself to a more mature level.

Borrowing From Buyer Behavior Theory

Historically, we have borrowed a great deal from other disciplines to build buyer behavior theory. However, I think within a decade, it is very likely that other disciplines will be actively interested in buyer behavior and consequently borrow from it a set of concepts and research tools. Implicit in this prediction is my conviction that we either already have or will very soon have richness of thinking, variety of methodology and respectability of the discipline to motivate other disciplines to search for relevant concepts and methods from buyer behavior. It is interesting,

therefore, to speculate which disciplines are likely to borrow what from buyer behavior theory. I have described below three types of borrowing activities: (1) less mature disciplines of social science borrowing the methodology of research in buyer behavior, (2) older social sciences borrowing concepts and theories of buyer behavior, and (3) hard sciences borrowing both the theory and methodology of buyer behavior discipline.

Less Mature Social Sciences

Relative to some mature social sciences such as macroeconomics and experimental psychology, the buyer behavior discipline looks less mature. By the same token, there are many other social sciences which are even less mature than buyer behavior. I include political science, parts of sociology, history, religion, home economics, law, and public health in the category of less mature disciplines in social sciences. Just as we have borrowed from psychology and economics, I believe these disciplines are likely to borrow from us. In fact, this is already evident from the recent trend of citing marketing and buyer behavior references in these disciplines.

The less mature social sciences are likely to borrow the research methods identified and routinely utilized in buyer behavior. This includes longitudinal panels, cost-oriented sampling procedures, the survey methods of data collection, and the use of multivariate methods. In addition, there is always the possibility of utilizing marketing strategies and tactics to diffuse radical innovations in each of those disciplines.

Mature Social

Some of my colleagues may not agree with me but, I think many of the traditional social sciences to which we owe so much are likely to at first participate in understanding of buyer behavior, and eventually to borrow from it. My prediction is based upon two facts. First, these traditional disciplines are currently facing the crisis of relevance because foundations of their theory and research have been based on less critical areas of human behavior. I include experimental psychology, social psychology and small group theory among others who have encountered this crisis in recent years. Second, many of the traditional disciplines have built formal models of behavior which have tended to be unrealistic or have become obsolete due to unprecedented technological change in our society in the last thirty years. I include the utility theory, micro theory of the firm and allied areas of economics and decision making as illustrative of this type of social sciences. In their search for societally relevant and useful problem areas, and to build realistic theories to help solve them the probability is extremely high that buyer behavior will become center of attention because social problems of a mass consumption society tend to be directly reflected in it.

Due to the problem-solving interests of these traditional social

sciences, I believe the traditional disciplines are likely to be more interested in the theoretical concepts and substantive findings than in the research methodology of buyer behavior. This looks also plausible in view of the fact that the traditional disciplines tend to be rich in methodology. I think buyer behavior has a lot to offer to the utility theory in economics from its thinking on choice behavior anchored to cognitive-evaluative structures. Similarly, I think we have a lot to offer to social psychology both in theory and research methodology in terms of conducting complex longitudinal studies of attitude change and brand choice behavior in naturalistic settings. Third, the growing literature and theory on diffusion of innovations in buyer behavior may well enable rural sociology to rethink diffusion theory. Finally, the recent emergence of efforts to build test batteries to measure life styles in buyer behavior is likely to significantly alter the thrust of personality tests in clinical psychology.

Although, most of the borrowing from buyer behavior by the traditional social sciences is likely to be with respect to concepts and substantive findings, there are certain areas of research methodology which may also be useful to them. These include the survey research aspects entailed in the design and execution of large scale studies in naturalistic settings.

Hard Sciences

By hard sciences, I mean natural sciences and engineering based on physics, mechanics, chemistry and biochemistry. The hard sciences have reached a level of maturity in their own disciplines to an extent whereby it is inevitable for them to broaden their horizons. I expect them, therefore, to concern themselves with the problem of social consequences arising from technology and depletion of natural resources. This includes, for example, areas of pollution of resources, urban planning, and the like. Recently, we have witnessed research undertakings by the hard sciences in those areas of social concern which typically have been the domain of social scientists. It seems inevitable, therefore, that sooner or later, the hard sciences are likely to be exposed to, and interested in buyer behavior. When that happens, it is equally inevitable that they will extensively borrow both the substantive findings and research methodology, because the newer research areas will force them to examine alternative theories and methodology. For it is generally conceded that the concepts of hard sciences may only be analogously related to social problems.

Conclusions and Discussions

In this paper, I have attempted to speculate on the future of buyer behavior theory. My speculations were limited to forecasting major directions which it is likely to take in the coming decade. These

predictions were based on the reviewing of historical perspective of buyer behavior theory and taking into account the contemporary events. I have suggested that four major changes are likely to occur in the structure of buyer behavior theory. They are (1) development of criteria to evaluate theories of buyer behavior, (2) construction of standardized tests and scales to measure buyer behavior constructs, (3) complex model building in an inductive manner with the use of several statistical procedures, and (4) research emphasis on nonpurposeful behavior. Second, I have suggested three major ways by which buyer behavior theory is likely to broaden its horizons, namely (1) development of cross-cultural theories of buyer behavior, (2) research and theories of buyer behavior for public policy purposes, and (3) research on social and environmental problems created by mass consumption societies. Third, I have predicted that a number of other disciplines will actively engage in buyer behavior and, therefore, substantially borrow research methodology and theory typically identified with buyer behavior theory. Specifically, I have suggested that (1) less mature social sciences such as political science, law, education and public health will borrow research methodology, (2) more mature and older social sciences are likely to borrow the concepts and theories from buyer behavior in their efforts to become more relevant and realistic disciplines, and (3) some natural sciences will borrow both methodology and theory from buyer behavior in the process of broadening their horizons to understand social consequences of technology.

Despite the bright predictions for the future of buyer behavior theory, I think there are some identifiable ailments in today's theories of buyer behavior which may impede the achievement of these predictions. First, most theories look upon buyer behavior as the consequence of some form of the decision-making process, and thus implicitly concede that buyer behavior consists of only goal-directed behavior. This may very well restrict the horizons to which buyer behavior theory can be broadened. Second, a large number of theories of buyer behavior often examine the buyer decision process from the point of view of marketing. While marketing management has made the greatest use of findings and concepts of buyer behavior, there is no reason why others from different viewpoints cannot utilize the same concepts and findings. Not only has this tendency made buyer behavior theory somewhat myopic, it has produced a terminology and vocabulary for buyer behavior which impedes its extension to nontraditional areas. Third, it seems that most theories of buyer behavior tend to overemphasize the process leading up to behavior and underemphasize the buying behavior or the antecedent and subsequent events which surround the behavior. Unless we consciously strive to remove these ailments, buyer behavior theory may take longer time to gain respectability across disciplines.

Footnotes

1. Jagdish N. Sheth is Professor of Business and Research Professor at the University of Illinois

References

- Andreasen, A. R. Attitudes and customer behavior; a decision model. L. E. Preston (ed.) New Research in Marketing, Berkeley: University of California, 1965, 1-16.
- Bass, F. M. et. al. Mathematical Models and Methods in Marketing, Homewood, Illinois: Irwin, 1961.
- Burk, M. C. Survey of interpretations of consumer behavior by social scientists in the postwar period. Journal of Farm Economics, 1967, 49, 1-31.
- Engel, J. F., Kollat, D. T., and Blackwell, R. D. Consumer Behavior, New York: Holt, Rinehart & Winston, 1968.
- Guest, L. Consumer analysis, Annual Review of Psychology, 1962, 13, 315-344.
- Hansen, F. Consumer Choice Behavior, New York, Free Press, 1972.
- Howard, J. A. Marketing Theory, Boston: Allyn and Bacon, 1965.
- Howard, J. A., Sheth, J. N. The Theory of Buyer Behavior, New York: Wiley, 1969.
- Katona, G. Rational behavior and economic behavior, Psychological Bulletin, 1953, 60, 307-318.
- Massy, W. F., Montgomery, D. B., and Morrison, D. G. Stochastic Models of Buying Behavior, Cambridge: MIT Press, 1970.
- Nicosia, F. M. Consumer Decision Processes, Englewood Cliffs: Prentice-Hall, 1966.
- Perloff, R. Consumer analysis, Annual Review of Psychology, 1968, 19.
- Sheth, J. N. A review of buyer behavior, Management Science, 1967, 13, B718-B756.
- Sheth, J. N. A theory of family buying decisions, P. Pellemans (ed.) Insights in Consumer and Market Behavior, Namur: Namur University, 1971, 32-49.
- Sheth, J. N., A theory of industrial buying decisions, Faculty Working Paper No. 61, Urbana: College of Commerce, University of Illinois, 1972.
- Twedt, D. W. Consumer psychology, Annual Review of Psychology, 1965, 16, 265-294.
- Venkatesan, M. Novelty seeking, Working Paper, University of Iowa, 1972.
- Zaltman, G., Pinson, C. and Angelmar, R. Metatheory and Consumer Behavior Research, New York: Holt, Rinehart, and Winston (in press).

THE CURRENT STATUS OF CONSUMER BEHAVIOR RESEARCH:
DEVELOPMENTS DURING THE 1968-1972 PERIOD¹

David T. Kollat, Management Horizons, Inc.
Roger D. Blackwell, The Ohio State University
James F. Engel, Wheaton College²

While preparing the first edition of Consumer Behavior (Engel, Kollat, Blackwell, 1968) it became apparent that a number of basic problems had retarded progress in consumer research prior to 1968 and were likely to continue to do so unless they were recognized and resolved, or at least accommodated. These problems were concerned with the development of what was termed a research tradition or strategy of inquiry, and were later published in an article appearing in the Journal of Marketing Research (Kollat, Engel, Blackwell, 1970).

During the last twelve months the authors have been revising the Consumer Behavior text and have therefore had the opportunity to review consumer research published during the 1968-1972 period. This proved to be a formidable challenge since the volume of research during this period was nearly as large as the total body of knowledge that existed at the beginning of the period.

The purpose of this paper is to evaluate this voluminous output in terms of the progress that has been made toward the development of a research tradition in consumer behavior. The criteria that are used are the same as those that were employed in the earlier JMR discussion. Since some of the analysis is critical, it should be remembered that many evaluations are generalizations having notable exceptions, and that the intent is to try to increase the effectiveness and efficiency of future efforts rather than critique the efforts of individuals.

Perspectives on Progress

Utilization of Consumer Behavior Constructs Theories and Models

Prior to 1968, the majority of consumer research utilized, explicitly or implicitly, hypothetical constructs, theories, and what Nicosia (1966) has called "reduced-form" models. Examples include motivation, perception, learning, personality, attitudes and attitude change, social class, and risk-taking. These constructs have been employed in a variety of ways in an attempt to explain and/or predict some aspect of consumer behavior.

While these constructs are often significant and useful, there have been many instances where they have not been used properly. As Jacoby (1969) and others have shown, this problem is evident in many investigations of the relationship between personality and consumer behavior.

Jacoby illustrates his points by re-analyzing Evans' (1959) study of the personality differences between Ford and Chevrolet owners. Evans employed the Edwards Personal Preference Schedule and found only one difference significant at the .05 level. From this, he concluded that personality had little, if any, relationship to consumer behavior. Jacoby shows that an entirely different picture emerges when the data are re-examined using specific hypotheses derived from a conceptual-psychological orientation. Specifically, Jacoby's analysis of Evans' data yielded 8 out of 11 correct predictions.

Unfortunately, the Jacoby example is not an isolated situation. For example, Brody and Cunningham (1968); Engel, Kollat, and Blackwell (1969); Kassarian (1971); and Wells and Tigert (1971) have demonstrated the value of personality and personality-oriented variables when used properly. In another area, there has been heated debate for many years about the relationship between attitudes and behavior. Yet Sheth and Talaryzk (1972) and Bass and Talaryzk (1972) have shown the importance of attitude in predicting consumer behavior when the theoretical basis for the concept is understood and used properly.

Thus, developments in this area during the last five years are mixed. On the one hand there is evidence that in many instances constructs and theories have been used improperly. On the other hand, there are an encouraging number of instances where proper utilization has yielded useful insights and results. Both developments suggest that when an aspect of consumer behavior is studied it is necessary to make certain that theoretical aspects are examined in detail and the theory is used correctly.

The second problem in using hypothetical constructs, traditional theories, and reduced-form models is that each plays a limited role in that consumer behavior is influenced by a variety of phenomena interacting in complex ways. According to Nicosia (1969), James Morgan (1958) was probably the first to point out that the exponential growth in the number of determinants of consumer behavior was causing increasing perplexity. During the last five years this problem has intensified to the point where it has become almost unmanageable.

Since 1966, several attempts have been made to design models comprehensive enough to deal with the multiplicity of determinants and the interrelationships of various constructs, theories, and reduced-form models. Nicosia (1966), Howard and Sheth (1969), Engel, Kollat and Blackwell (1968), and many others (Andreason, 1965; Ehrenberg, 1969) have proposed models, or frameworks, or conceptual schemes, of varying degrees of comprehensiveness and sophistication.

These models have had little influence on consumer research during the last five years. Indeed, it is rare to find a published study that has utilized, been based on, or even influenced by, any of the models identified above (Pellemans, 1971; Dominquez and Burger, 1971).

In 1968 the authors discussed the "conceptualization artifact" problem. This refers to the fact that without an integrative model the researcher does not know what variables should be included and controlled. Therefore, it is impossible to determine how many of the "significant" and "nonsignificant" findings would change if these variables were not neglected.

As mentioned above, the consumer behavior literature has doubled during the last five years. This constitutes a remarkable achievement by almost any standard. Unfortunately, however, it would not be surprising if over 90 percent of the findings and lack of findings prove to be wrong because they are artifacts of the reduced-form conceptualizations that have been used.

Why do researchers avoid the use of integrative models? One of the most commonly mentioned reasons is that the models are wrong; they are overly-simplistic or merely frameworks that have some expositional value, at least in the constructors' mind(s). While these points of view are plausible, and perhaps true, they cannot be accepted as true at this point in time because they have not been subjected to adequate empirical testing. In our review of the literature we have not found a single test of the Nicosia or Engel, Kollat and Blackwell models. Zaltman, Pinson and Angelmar (in press) arrived at the same conclusion.

Farley and Ring's (1970) empirical test of the Howard-Sheth model is perhaps the most interesting development in this area during recent years.

Their work indicates that considerable advances are required in measures before valid assessment of that model can be established. Their research also demonstrates that the Howard-Sheth model is more than an impressive flowchart; rather, the implied relationships in the model provide some directions for deriving inferences and testing relationships.

Farley and Ring's efforts suggest another reason why integrative models have not been tested -- some, including our own are partially or totally untestable. This suggests the need to devote more attention to the modeling-testing sequence. Most comprehensive models attempt to do at least two things. First, they record which variables are known to interact with which other variables. Second, they reveal which interactions need to be and have not yet been studied (theoretically and empirically). Most of the research used to construct these models proceeds from theoretical statements -- "Y is caused by X, Z and V" -- to direct empirical tests.

As Nicosia (1969) has pointed out, one of the shortcomings of this procedure is that the statistics obtained may actually be produced by different causal networks of interactions among the variables. As long as the interpretation of statistical results is ambiguous, it is not clear which theory is actually being tested; thus, the empirical results cannot be used to refine the original idea of how the phenomenon works. Rather, the result is an endless cascade of qualifications, and an unmanageable number of empirically-tested and non-rejected hypotheses.

To overcome these problems Nicosia recommends the insertion of a methodological operation between the set of theoretical statements and the empirical test. Predictably, the operation is a formal mathematical specification of the network of interactions the researcher has in mind. This is done by translating the hypothesized network (flowcharts) into formal models.

During the last five years there has been little work done in building, analyzing, and testing sophisticated mathematical models that predict and explain brand choice on the basis of interactions among a variety of variables over time. In the past, those knowledgeable in substantive areas have typically lacked super-sophisticated mathematical skills. Simultaneously, individuals possessing modeling expertise have often lacked a rigorous understanding of the substantive dimensions of the behavior being modeled. There is encouraging evidence that some terminal degree programs are correcting this skill imbalance. Hopefully a merger of these two types of competences will accelerate progress in the future.

Research Priorities

Prior to 1968, most consumer research occurred because of the availability of data, the convenience of research and mathematical techniques, and/or the attractiveness and appeal of certain behavioral constructs. While this orientation was understandable and justifiable in the short run, it is not effective, efficient, or responsible in the long run. Accordingly, the authors recommended that research priorities be established. These priorities should identify what "aspects" of consumer behavior are of the greatest importance, and what phenomena need to be investigated so that these key areas can be understood.

There has not been any progress in these areas during the 1968-1972 period. Most research -- perhaps as much as 95 percent -- continues to be data-technique-construct motivated and oriented. Thus, as never before, the profession faces what Robertson and Ward (1972) have recently called the "payoff dilemma":

The crux of the consumer behavior researcher's dilemma is that he seeks to engage in theoretical and conceptual research

and, at the same time to meet the "action" needs of its users. This is a familiar dilemma for the researcher in an applied field and places him in a difficult and sometimes compromising role. Judging from the results to date he has not handled the role conflict very well, for the consumer behavior field is both immature theoretically, yet has failed for the most part to meet the needs of its users -- marketing management, government and consumerist advocates, and most of all consumers.

The role conflict described by Robertson and Ward is one of the most serious problems facing the profession. It is rare to find a research study that discusses the role and implications of that study in terms of the development or testing of a theoretical understanding of consumer behavior. Alternatively, many studies that purport to be user oriented either avoid implications entirely or discuss them at an embarrassingly superficial level.

It is interesting to speculate about what the current status of the space program would be if NASA had used our ad hoc approach to selecting research topics. Some people feel that the present arrangement is as it should be; that researchers should be free to research whatever they want to regardless of its relevance. Curiously, many of the same people also talk about the social responsibilities of corporations. What are the social and ethical responsibilities of consumer researchers? How do we justify the resources we are consuming? Admittedly the field is young, but does the totality of individual efforts constitute a logical and defensible program that will yield progress?

Use of Longitudinal and Experimental Designs

Consumer behavior researchers typically use three types of research designs -- cross-sectional surveys, longitudinal, and experimental or quasi-experimental. Historically, cross-sectional surveys have been the most common.

The appropriateness of each method depends, of course, on the type of problem, the reasons for the study, the research budget, and the researcher's conceptualization of the problem. On balance, however, longitudinal and experimental designs are preferable, and, hence, in 1968 more widespread use of these two designs was recommended.

Mixed progress has been made in this area during the 1968-1972 period. Although longitudinal studies are still relatively rare (Pennington, 1968), the growing use of experimental designs is encouraging (Granger and Billson, 1972; McConnell, 1968).

Standardized Definitions

The pre-1968 consumer behavior literature was replete with varying definitions of what were presumably the same variables and constructs. Brand loyalty, opinion leaders, innovation, culture, personality, information seeking, and impulse purchasing are but a few examples of constructs that have a bewildering array of definitions.

Quite obviously, the definition heterogeneity problem makes it difficult and hazardous to compare, synthesize, and accumulate findings. Accordingly, the authors recommended the development of standardized definitions, or at least agreed upon points of departure.

Unfortunately, no progress has been made in this area during the last five years. Rather, in many cases, there has been a proliferation of

definitions. Consequently, although there is considerably more research output, progress is not as great as it would have been if standardized definitions had been developed and adopted.

Standardized Variable Categories

In 1968, the authors also pointed out the category heterogeneity problem that plagues many variables and constructs. Family life cycle, family role structure, and social class are examples of constructs that suffer from this problem.

Since this category variance also makes it difficult to compare and synthesize findings, it was suggested that standardized variable categories be developed and adopted. Instead of improving, this problem has intensified during the last five years. It is difficult to understand why it is possible to develop standardized demographic variable categories but not standardized behavioral categories.

Richer Dependent Variables

Regardless of the complexity of the dependent variable, researchers typically measure it unidimensionally. If dependent variables were measured multidimensionally, the independent variables that are significant might change. Hence, wider use of multidimensional measures was recommended in 1968.

There have been many attempts to deal with the "dimensionality artifact" problem during the 1968-1972 period. The growing use of multidimensional scaling techniques is perhaps the most encouraging development in this area (Green and Rao, 1972; Day, 1972; Lehmann, 1972). There have been other interesting developments. For example, studies of brand loyalty (measured unidimensionally) have been characterized by the absence of significant relationships with consumer characteristics. Yet, when loyalty is measured multidimensionally, significant relationships have surfaced (Carman, 1970; Day, 1969). Overall, important progress has been made in the use of multidimensional measures of dependent variables.

Replication

Prior to 1968, replication was rarely practiced in consumer research. Unfortunately, this behavior has not changed during the last five years. Thus, most findings and propositions continue to be based upon single studies by a single researcher, or multiple studies by multiple researchers using different definitions and variable categories. Quite obviously, this practice invites invalid conclusions due to unusual sample characteristics, distortion in experimental control, and a variety of other methodological artifacts. Hopefully, a replication tradition will be promoted and practiced in the future.

There have been encouraging developments in a related area. Prior to 1968, the literature was dominated by "one-shot" studies using samples that were small, out-of-date, or questionable on other grounds; e.g., college students, women's club members, etc. The one-shot characteristic limits the researcher's ability to investigate the phenomenon in a rigorous and comprehensive manner.

In contrast, research efforts that seem to have had a decisive impact on the discipline during the last five years are major research programs. These programs are longer-term efforts that systematically investigate many dimensions of a phenomenon, and usually, but not always, involve larger and better quality samples. Examples include the King-Summers (1970) thrust in

opinion leadership and diffusion, the Pessemier-Tigert-Wells (Wells and Tigert, 1971) investigations of psychographics and other profiling and clustering techniques, the Green (Green and Rao, 1972) stream in multi-dimensional scaling, and the Howard-Sheth, et.al. (Howard, 1971) tests of their theory of buyer behavior.

These and other researchers have attained critical mass in studying some aspect of consumer behavior. If the research program approach were used in other areas -- social class, reference groups, purchase intentions, store choice, and shopper profiles, to mention just a few -- progress would probably accelerate.

Generalizing Findings

To what extent are consumer behavior findings artifacts of the research design, subjects used, and variables controlled? To what degree are findings derived from an analysis of a specific type of consumer decision applicable to other types of decisions? Certainly in many cases it is not proper to generalize findings across decision situations, or from a research design to the "real world". On the other hand, generalizing as far as possible avoids researching consumer behavior in unnecessarily minute detail. For these reasons, the authors have pointed up the growing need for classification systems for types of decisions and choices which, if properly designed, would permit a legitimate degree of generalization.

During recent years, these issues have received more attention. For example, Farley, Howard and Weinstein (1971) have analyzed the stability of attitude structures. Similarly, Pessemier and Bruno (1971) have studied the stability and reliability of activity and attitude measures. In both instances, important insights were uncovered.

Simultaneously, there have been numerous attempts to develop various types of classification schemes. Some efforts have focused primarily on classification techniques such as factor or cluster analytic models and latent structure analysis (Myers and Nicosia, 1969). Others have experimented with a variety of behavioral concepts (Montgomery and Silk, 1971; Anderson, 1971; Johnson, 1971).

From a theoretical perspective, Robertson and Ward (1972) have encouraged the development of more "middle-range" theories. These are theories intermediate to the minor working hypotheses evolved in abundance during the day-to-day routines of research, and the all-inclusive speculations comprising a master conceptual scheme.

Although these efforts have not resolved the questions of generalizability, at least attempts are being made to confront and understand the issues. Future efforts using alternative conceptual schemes, empirically derived classifications, and new analytic techniques should accelerate progress.

Broadening the Uses and Horizons of Consumer Research

As mentioned earlier, there are three major end users of consumer research -- marketing management, government and consumerist advocates, and consumers. User-oriented research has continued to be directed primarily toward marketing management. Some progress has been made in beginning to understand the behavioral problems of the disadvantaged, but very little research attention has been devoted to consumer exploitation or protection, or the behavioral implications of legal and regulatory actions. Hopefully this imbalance will be corrected in the future.

During the 1968-1972 period consumer research focused almost exclusively on the micro behavior of individual consumers or market segments, and ignored macro behavior issues. As Nicosia and Glock (1968) have pointed out, many economic and social problems cannot be solved unless we gain an understanding of the relationships between changing patterns of consumption and changes in social and cultural values. Consider these issues:

1. What effect would smaller families have on consumption patterns?
2. Will the new values and life styles of some youth change as they grow older? If not, how will they affect consumption?
3. Will the women's liberation movement become more important? If so, how will it affect purchasing and consumption patterns?
4. How will the cost of clean air and water change consumption patterns?
5. How will the four-day work week affect expenditure patterns?

These are not idle issues. Some observers estimate that variations in these behavioral patterns could make a difference of \$500 billion in the 1980 Gross National Product (Silberman, 1971). These types of questions deserve serious attention.

Information Summary and Retrieval Systems

Several years ago it became clear that there would be a research explosion and that it would be difficult for both researchers and practitioners to have an awareness and working knowledge of published research relevant to their problems unless steps were taken to accommodate the problem. Two steps were recommended: annual literature reviews, and the development of a consumer behavior research retrieval system.

To date neither of these steps have been taken. Unless action is initiated in the near future, by 1980 it may be impossible to have a comprehensive understanding of even relatively small components of consumer behavior.

Summary and Conclusions

Risking oversimplification, Exhibit 1 summarizes the authors' evaluations of developments in consumer research during the 1968-1972 period. Some will feel that the evaluations are too rigorous, while others will contend that they are not strict enough. Inevitably, important contributions have been omitted although certainly not intentionally.

Good progress has been made in some areas -- the use of experimental designs, the use of richer dependent variables, and the development of research programs rather than "one-shot" studies. However, in most areas progress has been limited or nonexistent.

Recently the National Science Council (1969) evaluated the state of knowledge in the behavioral sciences. They concluded that actual accomplishment has not been consistent with the magnitude of effort. Unfortunately, this conclusion appears applicable to the field of consumer research. Programs designed to resolve or at least accommodate the issues discussed above might help correct the results-effort imbalance in the future.

Exhibit 1

Summary of Developments in Consumer Research
During the 1968-1972 Period

Area	Evaluation
Utilization of consumer behavior constructs, theories, and models	Limited progress
Testing of consumer behavior models	Limited progress
Development of research priorities	No progress
Use of longitudinal designs	Limited progress
Use of experimental designs	Good progress
Development of standardized definitions	No progress
Development of standardized variable categories	No progress
Utilization of richer dependent variables	Good progress
Development of a replication tradition	Limited progress
Development of research programs rather than "one-shot" studies	Good progress
Generalizing findings	Fair progress
Broadening the uses and horizons of consumer research	Limited progress
Developing information summary and retrieval systems	No progress

Footnotes

1. Adapted from Engel, Kollat and Blackwell (in press), Chapter 27.
2. David T. Kollat is Vice President-Research, Management Horizons, Inc., Columbus, Ohio. He was formerly Professor of Marketing, The Ohio State University. James F. Engel is Professor of Communications, Graduate School, Wheaton College. He was formerly Professor of Marketing, The Ohio State University. Roger D. Blackwell is Associate Professor of Marketing, The Ohio State University and Vice President, Management Horizons, Inc.

References

- Anderson, T. W., Jr. Identifying the convenience oriented consumer. Journal of Marketing Research, 1971, 8, 179-183.
- Andreason, A.A. Attitudes and consumer behavior: a decision model. In L.R. Preston (Ed.), New research in marketing. Institute of Business and Economic Research: University of California, Berkeley, 1965, 1-16.
- Bass, F.M. & Talaryzk, W.W. An attitude model for the study of brand preference. Journal of Marketing Research, 1972, 9, 93-96.
- Brody, R.P. & Cunningham, S.M. Personality variables and the consumer decision process. Journal of Marketing Research, 1968, 5, 50-57.
- Carman, J.R. Correlates of brand loyalty: some positive results. Journal of Marketing Research, 1970, 7, 67-76.
- Day, G.S. A two-dimensional concept of brand loyalty. Journal of Advertising Research, 1969, 2, 29-35.
- Day, G.S. Evaluating models of attitude structure. Journal of Marketing Research, 1972, 9, 279-286.
- Dominquez, L.V. & Burger, P.C. An empirical analysis of the process of consumer behavior. In F.C. Allvine (Ed.), Relevance in marketing. Chicago: American Marketing Association, 1971, 391-396.
- Ehrenberg, A.S.C. Toward an integrated theory of consumer behavior. Journal of the Marketing Research Society, 1969, 11, 305-337.
- Engel, J.F., Kollat, D.T. & Blackwell, R.D. Consumer behavior. New York: Holt, Rinehart & Winston, 1968.
- Engel, J.F., Kollat, D.T. & Blackwell, R.D. Consumer behavior. 2nd Edition. New York: Holt, Rinehart & Winston, in press.
- Engel, J.F., Kollat, D.T. & Blackwell, R.D. Personality measures and market segmentation. Business Horizons, 1969, 12, 61-70.
- Evans, F.B. Psychological and objective factors in the prediction of brand choice: Ford versus Chevrolet. Journal of Business, 1959, 32, 340-369.
- Farley, J.U. & Ring, L.W. An empirical test of the Howard-Sheth model of buyer behavior. Journal of Marketing Research, 1970, 7, 427-438.
- Farley, J.U., Howard, J.A. & Weinstein, D. An investigation of stability in attitude structure toward a product class. In F.C. Allvine, (Ed.) Relevance in marketing. Chicago: American Marketing Association, 1971, 337-345.
- Granger, C.W.J. & Billson, A. Consumers attitudes toward package size and price. Journal of Marketing Research, 1972, 9, 239-248.
- Green, P.E. & Rao, V.R. Applied multidimensional scaling. New York: Holt, Rinehart & Winston, 1972.
- Howard, J.A. & Sheth, J.N. The theory of buyer behavior. New York: John Wiley & Sons, 1969.
- Howard, J.A. New directions in buyer behavior research. In F.C. Allvine, (Ed.), Relevance in marketing. Chicago: American Marketing Association, 1971, 375-380.
- Jacoby, J. Towards defining consumer psychology: one psychologist's views. Paper presented at the American Psychological Association 77th Annual Convention, September, 1969.
- Johnson, R.A. Market segmentation: a strategic management tool. Journal of Marketing Research, 1971, 8, 13-18.
- Kassarjian, H.J. Personality and consumer behavior: a review. Journal of Marketing Research, 1971, 8, 409-418.
- King, C.W. & Summers, J.O. Overlap of opinion leadership across consumer product categories. Journal of Marketing Research, 1970, 7, 43-50.
- Kollat, D.T., Engel, J.F. & Blackwell, R.D. Current problems in consumer behavior research. Journal of Marketing Research, 1970, 7, 327-332.

- Lehmann, D.R. Judged similarity and brand switching data as similarity measures. Journal of Marketing Research, 1972, 9, 331-334.
- McConnell, J.D. The price-quality relationship in an experimental setting. Journal of Marketing Research, 1968, 5, 300-303.
- Montgomery, D.B. & Silk, A.J. Clusters of consumer interests and opinion leaders' spheres of influence. Journal of Marketing Research, 1971, 8, 317-321.
- Morgan, J.M. A review of recent research on consumer behavior. In L. Clark (Ed.) Consumer behavior. New York: Harper & Row, 1958, 93-108.
- Myers, J.G. & Nicosia, F.M. On the dimensionality question in latent structure analysis. In P.R. McDonald (Ed.) Marketing involvement in society and the economy. Chicago: American Marketing Association, 1969, 145-149.
- National Science Council, The behavioral and social sciences. Englewood Cliffs, New Jersey: Prentice-Hall, 1969.
- Nicosia, F.M. Consumer decision processes. Englewood Cliffs, New Jersey: Prentice-Hall, 1966.
- Nicosia, F.M. & Glock, C.Y. Marketing and affluence: a research prospectus. In R.L. King (Ed.) Marketing and the new science of planning. Chicago: American Marketing Association, 1968, 510-527.
- Nicosia, F.M. Brand choice: toward behavioral-behavioristic models. Paper presented to the Symposium on Behavioral Sciences and Management Sciences in Marketing, Chicago, June, 1969.
- Pellemans, P.A. Relationships between attitude and purchase intention toward the brand. Belgium: Publication Universitaires, Namur University, 1971.
- Pennington, A.L. Customer-salesman bargaining behavior in retail transactions. Journal of Marketing Research, 1968, 5, 255-262.
- Pessemier, E. & Bruno, A. An empirical investigation of the reliability and stability of selected activity and attitude measures. In D.M. Gardner (Ed.) Proceedings, Association for Consumer Research, 1971, 389-403.
- Robertson, T.S. & Ward, S. Toward the development of consumer behavior theory. Paper presented at the Fall Conference of the American Marketing Association, August, 1972.
- Sheth, J.N. & Talaryzk, W.W. Perceived instrumentality and value importance as determinants of attitudes. Journal of Marketing Research, 1972, 9, 6-9.
- Silberman, C.E. The U.S. economy in an age of uncertainty. Fortune, January, 1971, 72 ff.
- Wells, W.D. & Tigert D.J. Activities, interests and opinions. Journal of Advertising Research, 1971, 11, 27-35.
- Zaltman, G., Pinson, C.R.A. & Angelmar, R. Metatheory and consumer research. New York: Holt, Rinehart & Winston, in press.

AN EXAMINATION OF CONCEPT VALIDITY

Reinhard Angelmar, Gerald Zaltman
and Christian Pinson¹
Northwestern University

Introduction

This paper will examine various types of concept validity of relevance to the study of consumer behavior. The issue of concept validity may well be the Achilles heel in the study of consumer behavior. This statement itself has some consensual validity. For example, in an article devoted to the problems in consumer behavior research, Kollat, et. al. (1970:328) pointed out that "future progress in consumer behavior research will depend on overcoming several problems with commonly used variables and constructs." Kassarjian (1971), in his review of personality concepts in consumer behavior, has similarly emphasized the crucial role of having valid concepts.

Any discussion of the validity of concepts involves metatheoretical considerations. Metatheory is the investigation, analysis, and the description of (1) the technology of building theory, (2) the theory itself, and (3) the utilization of theory. Concepts are the essential building blocks of theory and a theory can be no better than its concepts. Especially relevant is the issue of the validity of concepts and hence the validity of theories.* The goal of this paper is to contribute to the solution of the present conceptual problems by reviewing several commonly used but rarely made explicit types and criteria of concept validity.

Types of Concept Validity

Observational Validity

Seven types of concept validity will be discussed below. These are shown in summary form in Table 1. The first and most traditional approach considers only observational concepts to be valid. In its extreme form this approach, which has been called operationism, requires the exhaustive reducibility of all concepts to observations (Nagel, 1961). Hempel (1966:88) notes that the "central idea of operationism is that the meaning of every scientific term must be specifiable by indicating a definite testing operation that provides the criterion for its application." Concepts that are provided with such criteria are said to be operationally defined. Bridgman (1927) adds a further specification to this point of view, stipulating that different operations characterize different concepts which ideally should be designated by different terms. This version of operationism has also been called "definitional" operationism (Campbell, 1969).

The basic motivation underlying operationism is to "emancipate science from any dependency on unverifiable 'metaphysical' commitments" (Nagel, 1961:119) This view has been criticized on several counts. One counterargument notes that

* For a more complete treatment of the various dimensions of concepts, see Chapter 2 in Gerald Zaltman, Christian Pinson and Reinhard Angelmar, Metatheory and Consumer Research (New York: Holt, Rinehart & Winston, 1973).

TABLE 1: TYPES OF CONCEPT VALIDITY

- | | |
|--------------------------------------|---|
| 1. <u>OBSERVATIONAL VALIDITY</u> | The degree to which a concept is reducible to observations. |
| 2. <u>CONTENT VALIDITY</u> | The degree to which an operationalization <u>represents</u> the concept about which generalizations are to be made. |
| 3. <u>CRITERION RELATED VALIDITY</u> | The degree to which the concept under consideration enables one to predict the value of some other concept which constitutes the criterion. |
| 3a. <u>Predictive Validity</u> | A subtype of criterion-related validity in which the criterion measure is separated in time from the predictor concept. |
| 3b. <u>Concurrent Validity</u> | A subtype of criterion-related validity in which the criterion and the predictor concepts are measured at the same time. |
| 4. <u>CONSTRUCT VALIDITY</u> | The extent to which an operationalization <u>measures</u> the concept which it purports to <u>measure</u> . |
| 4a. <u>Convergent Validity</u> | The degree to which two attempts to measure the same concept through maximally different methods are convergent. It is generally represented by the correlation between the two attempts. |
| 4b. <u>Discriminant Validity</u> | The extent to which a concept differs from other concepts. |
| 4c. <u>Nomological Validity</u> | The extent to which predictions based on the concept which an instrument purports to measure are confirmed. |
| 5. <u>SYSTEMIC VALIDITY</u> | The degree to which a concept enables the integration of previously unconnected concepts and/or the generation of a new conceptual system. |
| 6. <u>SEMANTIC VALIDITY</u> | The degree to which a concept has a uniform semantic usage. |
| 7. <u>CONTROL VALIDITY</u> | The degree to which a concept is manipulable and capable of influencing other variables of interest. |

"if explicit definitions of all theoretical terms by means of observables could be carried out, theories would be incapable of growth and therefore useless" (Hesse, 1967:406-407). Hempel argues that the definitional operationist maxim would lead to a proliferation of concepts "that would not only be practically unmanageable but theoretically endless, and this would defeat one of the principal purposes of science; namely, the attainment of a simple, systematically unified account of empirical phenomena" (Hempel, 1966:94). Campbell also advances a theoretical and a practical argument against definitional operationism. The theoretical argument says that any specific measurement reflects not a single parameter of a scientific theory but is "a joint function of many scientific laws." The practical argument says that the doctrine does not take into account the ongoing effort to improve measurement devices.

A less extreme form of operationism has been advanced by Carnap (1956). His requirement involves only partial reducibility of concepts to observations. If a concept is introduced into some scientific system one must be able to construct some proposition containing this new concept which, together with one or several other propositions containing only already tested terms, entails observation statements whose truth can be directly tested. This approach avoids the disadvantages of the extreme operationist position while still guaranteeing the empirical significance of concepts.

Most consumer behavior researchers seem to be well aware of this type of concept validity. In fact, one sometimes gets the impression that the observational validity of concepts is emphasized too much, and at the expense of other validity-types (such as construct validity).

The next three types of concept validity to be dealt with have received substantial attention by psychologists. Here, we refer to content, criterion-related-, and construct-validity.

Content Validity

Content validity refers to "the degree that the score or scale being used represents the concept about which generalizations are to be made" (Bohrnstedt, 1970:91). In order to assess content validity, it is necessary to define the universe. Only if this is done can the representativeness of the measure be evaluated. Consider the example of the concept of opinion leadership. In order to determine whether a certain opinion leadership measure has content validity, the classes of behavior to which the concept refers have to be defined. Following this, it is possible to determine whether the observations which the measure implies are representative of this universe.

Criterion--related Validity

Criterion-related validity is concerned with how well the concept enables one to predict the value of some other concept which constitutes the criterion. It is also called empirical or practical validity (Campbell, 1960). Criterion-validity consists of two subtypes, predictive and concurrent validity. In predictive validity, the criterion measure is separated in time from the predictor concept, while for concurrent validity both concepts are measured at the same time.

The distinctive characteristic of criterion-related validity is that, due to the "socially institutionalized and valued nature of the 'criterion,' it is taken as an immutable given" (Campbell, 1960:547). Purchase behavior, which fits Campbell's characterization as well as any other variable, is a frequent criterion in consumer behavior research.

A study by Axelrod (1968) provides an example of predictive validation. Axelrod was interested in finding a "measure that not only reflects the immediate effect of a stimulus on a consumer but also predicts his subsequent purchase behavior." For this purpose he developed ten measures, tested each one and concluded that two of the measures had the highest predictive validity as far as short-term trends in purchase behavior are concerned. The predictive validity of each measure was determined as the percentage of obtained market as compared to predicted market.

Many studies in consumer behavior research consist of the simultaneous collection of measures of many consumer characteristics. One or several of these characteristics--usually those related to purchase behavior--are then taken to be the criterion, and their presence or absence (or their value if the criterion is quantitative) is "predicted" with the help of the remaining variables. Such studies are typical examples of concurrent validation. The better a concept "predicts" the criterion, the greater its concurrent validity.

A study by Robertson and Kennedy (1969) illustrates concurrent validation. Data on a number of consumers were collected. One of the characteristics measured was possession of a small home appliance. This was taken to be the criterion. The remaining variables were used to predict the possession of the appliance. The main result of the study was that, from among the variables considered, venturesomeness and social mobility had the highest concurrent validity.

Construct Validity

Construct validity refers to the extent to which an operationalization measures the concept which it purports to measure. Following Campbell (1960), three types of construct validity can be distinguished: convergent, discriminant, and nomological validity. The first two types can be considered together under the label "trait validity."

The distinguishing characteristic of trait validity is that, in contrast to criterion-related validity, there is "no a priori defining criterion . . . available as a perfect measure or defining operation" (Campbell, 1960) against which to check a new measure. Instead, all of the measures are considered to be fallible.

Convergent validity refers to the degree to which two attempts to measure the same concept through maximally different methods are convergent. Discriminant validity refers to the extent to which the measure of a concept is related to measures of other concepts from which it is supposed to differ (Campbell and Fiske, 1959).

A recent study by Jacoby (1972) illustrates convergent and discriminant validation procedures. The concepts investigated were opinion leadership for clothing, for alcoholic beverages, and for LP records. While these three concepts are supposed to be different, they are also supposed to be related.

This makes the establishment of discriminant validity somewhat more difficult than if the concepts chosen had been independent from one another. The methods consisted of self-designation, sociometric choice, and the key-informant technique.

Convergent validity of each concept was measured by the correlation between the results of the three methods to measure the same concept. For example, the correlation between the results of the key-informant method and the sociometric method for measuring opinion leadership for LP records was calculated. All of these correlations turned out to be quite high.

In order to establish discriminant validity three conditions have to be satisfied (Campbell and Fiske, 1959): (1) The convergent validity for any concept should be higher than the correlation between any measure of that concept and a different concept measured by a different method. For example, the correlation between opinion leadership for clothing as measured by the key informant technique and by the sociometric technique, should be higher than the correlation between opinion leadership for clothing as measured by self-designation and opinion leadership for alcoholic beverages as measured by sociometric choice. (2) The convergent validity for any concept should be higher than the correlation of a concept with another concept, when both are measured by the same method. For example, the correlation between opinion leadership for alcoholic beverages measured by self-designation and as measured by the key-informant technique, should be higher than the correlation between the same type of opinion leadership and opinion leadership for clothing, with both concepts being measured by the sociometric technique. (3) The same pattern of interrelationships ought to obtain between the correlations of different concepts measured by the same method and as measured by different methods. For example, the ranking by magnitude of the correlations between opinion leadership for clothing, alcoholic beverages, and LP records, measured by the key-informant technique, should be the same as the ranking of the correlations between these concepts where each is measured by different methods. In Jacoby's study all of the three conditions of discriminant validity were fairly well satisfied.

Nomological validity refers to the extent to which predictions based on the concept which an instrument purports to measure are confirmed (Cronbach and Meehl, 1955). Another study of opinion leadership illustrates this type of validation procedure. Corey was interested in determining whether a particular type of technique was a valid measure of opinion leadership (Corey, 1971). At the time of the study, a body of knowledge concerning certain characteristics of opinion leaders had been accumulated. Corey reasoned that if his measure were in fact a valid measure of opinion leadership, the people classified as opinion leaders by it ought to have the characteristics indicated by the literature. This turned out to be true. Hence, he concluded that his instrument was a valid measure of opinion leadership.

Systemic Validity

The next approach to concept validity is concerned with a concept's systemic validity. This refers to the extent to which a concept can "establish relations among concepts and contribute thereby to systemicity (theoretical fertility)" (Bunge, 1957:133). For example, the concept of reciprocity can be used to explain the adoption or purchase of a standardized product in an interpersonal selling situation. The concept of reciprocity refers to feelings of obligation experienced by one party (e.g., a potential buyer) as he perceives another party (e.g., a salesman) investing or expending scarce resources on his behalf.

Semantic Validity

Another dimension of concept validity refers to the extent to which a concept has a uniform semantic usage (Marx, 1963). This can be called semantic validity. Kollat et. al., in their article earlier, mention several concepts whose semantic validity is notoriously low, among them brand loyalty, innovation, culture, and motive. This dimension of concept validity is of particular importance for the comparison, accumulation, and synthesis of findings, activities all of which are basic to paradigmatic research (Kuhn, 1962).

Control Validity

The last but not least important dimension to be considered here is a concept's control validity. This refers to the extent to which a concept is manipulatable and capable of influencing other variables of interest. The concepts which one finds in consumer behavior research range all the way from directly manipulatable concepts such as price to indirectly manipulatable concepts such as attitude to not at all manipulatable concepts such as birthorder (Kirchner, 1971). From the point of view of control validity concepts such as birth order have little validity. Intermediate concepts such as attitude have control validity to the extent to which antecedents of attitudes can be manipulated, and to the extent to which there is a substantial relationship between attitudes and purchase behavior. For example, Bauer (1966:8) has asserted that "the validity of our assessments of attitudes is solely the utility of the inferred concept for understanding, predicting and influencing the behavior of individuals."

Conclusion

This paper has noted rather briefly several types of concept validity which are relevant for the study of consumer behavior. Greater sensitivity to these validity criteria should produce more concrete and more useful concepts in marketing research. This in turn should lead to stronger theories in consumer behavior contexts. It is felt that the validity of concepts in current use in marketing and particularly in consumer behavior, leaves much to be desired. Evaluating concepts in terms of the types of concept validity presented here should strengthen this present state somewhat. Attention should be given to other possible types of concept validity.

Footnote

1. Gerald Zaltman is Associate Professor of Behavioral Science, Director of Research, Graduate School of Management and Faculty Associate, the Center for the Interdisciplinary Study of Science and Technology, Northwestern University. Reinhard Angelmar and Christian Pinson are currently completing their doctoral requirements in the Marketing Department, Graduate School of Management, Northwestern University.

References

- Axelrod, J. N. Attitude measures that predict purchase. Journal of Marketing Research, 1968, 8, 3-18.
- Bauer, R. Attitudes, verbal behavior and other behavior. In Adler, L. and Crespi, I. (eds.), Attitude research at sea. Chicago: The American Marketing Association, 1966, 3-14.
- Bohrnstedt, G. W. Reliability and validity assessment in attitude measurement. In Summers, G. F. (ed.), Attitude measurement. Chicago: Rand-McNally, 1970.
- Bridgeman, P. W. The logic of modern physics. New York: Macmillan, 1927.
- Bunge, M. Scientific research, I and II. New York: Springer-Verlag, Inc., 1967.
- Campbell, D. T. Definitional versus multiple operationalism. In Et. Al., 1969, 2, 14-17.
- Campbell, D. T. Prospective: artifact and control. In Rosenthal, R. and Rosnow, R. L. (eds.), Artifact in behavioral research. New York: Academic Press, 1969, 351-382.
- Campbell, D. T. Recommendations for APA test standards regarding construct, trait, or discriminant validity. American Psychologist, 1960, 15, 546-553.
- Campbell, D. T. and Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105.
- Carnap, R. The methodological character of theoretical concepts. Minnesota Studies in the Philosophy of Science, 1956, I.
- Corey, L. G. People who claim to be opinion leaders: identifying their characteristics by self-report. Journal of Marketing, October, 1971, 35, 48-53.
- Cronbach, L. J. and Meehl, P. E. Construct validity in psychological tests. Psychological Bulletin, 1955, 52, 281-302.
- Helmstadter, G. C. Principles of psychological measurement. New York: Appleton-Century-Crofts, 1964.
- Hempel, C. G. Philosophy of natural science. Englewood Cliffs, N.J.: Prentice-Hall, 1966.
- Hesse, M. Laws and theories. In Edwards, P. (ed.), Encyclopedia of philosophy. New York: Free Press, 1967, 406-407.
- Jacoby, J. Establishing the construct validity of opinion leadership. Purdue papers in consumer psychology, 1972, 121.
- Kassarjian, H. H. Personality and consumer behavior: a review. Journal of Marketing, 1971, 8, No.4, 409-419.

- Kirchner, D. F. Personal influence, ordinal position and purchasing behavior. ACR-Proceedings, 1971, 82-98.
- Kollat, D. T., Engel, J. F., and Blackwell, R. D. Current problems in consumer behavior research. Journal of Marketing Research, 1970, 7, 327-332.
- Kuhn, T.S. The structure of scientific revolutions, 2nd ed. Chicago: University of Chicago Press, 1962.
- Marx, M H. The dimension of operational clarity. In Marx, M H. (ed.), Theories in contemporary psychology. New York: Macmillan, 1963, 187-202.
- Nagel, E. The structure of science. New York: Harcourt, Brace, 1961.
- Robertson, T. S. and Kennedy, J. N. Prediction of consumer innovators: application of multiple discriminant analysis. Journal of Marketing Research, February, 1969, 5, 64-69.

A MODEL FOR RESEARCH PLANNING
IN CONSUMER BEHAVIOR

Richard W. Pollay¹
University of British Columbia

Despite impressive growth, it is arguable whether consumer behavior has yet established the research tradition and theoretical cohesiveness indicative of either a unique or a mature discipline. It seems we are still in the stage where, as a discipline, a struggle for a clear identity continues.

This position will need some explication, given the volume of literature being provided, and such is the purpose of the first part of this paper. It will be a necessity in this section, however, to be restricted to making some overview generalizations about the state of development of the discipline. It is not possible, no matter how desirable, to adequately review the substantive research in the field. That literature is simply too voluminous. That task will be left to the more ambitious, for it has been nearly five years since the last comprehensive review (Engel, et.al., 1968; Perloff, 1968; Sheth, 1967) and I have to agree with Perloff (1968) when he stated that the job of summarizing even three years of literature at an earlier stage of development of the "breathlessly changing field of consumer psychology is a task too Herculean for the reviewer to undertake" (p. 460). Thus this paper will not provide a summary of the state of knowledge, nor will it make any quasi-theoretical observations about the relationships between the knowledge in the various subareas, but rather will begin by a brief examination of consumer behavior as a science, examining its apparent productivity and progress.

The need for a comprehensive theory has been expressed in the literature repeatedly, and despite the admirable attempts to provide us with such a comprehensive theory, it retains its validity and deserves reiteration. The existence of a comprehensive theory is a necessary, although hardly sufficient, condition for the effective planning of research efforts, and it is this planning function that is the focus of this paper. Even in the absence of such a theory, however, it is possible to do some research planning. To put it another way, it is not inevitable that the development of the discipline be the haphazard result of fortune, idiosyncratic interests of the individuals involved, and methodological tractability. We can engage in a concerted effort to channel our energies into developing the kind of knowledge that may be most helpful in the development of the theory we seek.

To organize research activities meaningfully, it helps to have a means of readily comprehending the "state of the art". One such means is presented here: a topological conceptualization of accumulated knowledge. The full applicability of this model to consumer behavior research depends on the existence of a well articulated and accepted theoretical framework. This will be freely admitted; but the value of the discussion is not negated by the current lack of such a theory. The model precipitates some reorganization of thought, directing one's attention to the planning function in research programs. It also can be of obvious value in structuring research within the areas in which we have theoretical fragments.

Even if that were not the case, the discussion has value in the long run by preparing the way for effective research program planning once a more comprehensive theory starts to take shape. The analysis of the discipline's informational needs and the development of research strategies can, should, and

will be carried to more sophisticated levels than this introductory paper can even attempt. Effective research program planning can eliminate redundancy of effort, insure the production of aggregatable research results, and offer a fuller spectrum of information.

The State of the "Science"

Judging by the volume of output, the study of consumer behavior is a healthy science indeed. But this veritable flood of articles, speeches, working papers and research reports is in itself a mixed blessing and not necessarily indicative of any evolution in the discipline--except a growth in the number of practitioners. One of the negative consequences is the simple problem of information overload. Put more precisely, "The dilution of the meritorious by floods of triviality makes the recognition of true scientific value particularly difficult"² (Polanyi, 1958, p. 149). In addition, much of that published is research that was the "result of the availability of data, the convenience of research and mathematical techniques, and/or the appeal of certain behavioral constructs" (Kollatt, et.al., 1970, p. 328). But the ease of measurement is a poor criterion for selection of topics worthy of research. The exercise of measurement capabilities loses meaning when the existing theories are not highly directive toward objects of measurement. This is often insufficiently recognized, even, or perhaps especially, by those who call most vigorously for even more powerful measurement devices. As Sheth (1967, p. B719) notes, the "excursion into multivariate complexity presumes a formal understanding of the buyer". He might have explicitly added, that it was and is a presumption of very questionable validity.

None of the currently popular theories truly provide a key to adequate comprehension of the "data" being generated, although they may provide a mnemonic device for a cataloguing of such data. While the data bank has increased considerably, this has led to as much confusion as comprehension and we seem to witness more productivity than progress. Again, this results from the "paucity, and quite possibly the absence, of any theoretical under-pinnings which could give direction, meaning and permanence to substantive investigations" (Perloff, 1968, p. 437).

The status of marketing and consumer behavior as sciences is reflected in the very fact that the question is so very self-consciously and chronically addressed. That literature (Dawson, 1971; Halbert, 1964; Hunt, 1971; Lee, 1965; Robin, 1970; Sheth, 1967) of which this paper is guilty of being a part, has been disappointing and often has not attracted the best known scholars despite its importance. The asking of the question indicates our aspiration to be a science, but also our failure to realize that aspiration.

Kuhn (1962) discusses several behaviors which he holds to be indicative of something other than a normal science. In his terminology the following characteristics give corroborative evidence to the labelling of a discipline's activities as either "crisis science" or "preparadigmatic" activity: the proliferations of theories, or versions of theories; the expression of explicit discontent; the recourse to philosophy; the disagreement over the importance of "problems", a debate of fundamentals; a faster growth of complexity than accuracy; and professional insecurity among the scientists of the area. Since consumer research seems to possess nearly all of these attributes,³ one is led to the conclusion that the study of consumer behavior is not truly a science no matter how much we may utilize what we perceive to be scientific methods.⁴ In Kuhn's classification, consumer behavior is in a pre-paradigm state.

We are led once again to the conclusion that consumer behavior is sorely in need of theoretical improvements. This is hardly a novel conclusion and has been made in the literature repeatedly, perhaps best in the major reviews of the literature. Several other papers have discussed various aspects of the role of theory in consumer behavior and marketing. Halbert (1964) in an early paper discusses the requirement for a theory in marketing and describes especially well the requirements of syntactics, semantics and pragmatics. Hunt (1971) and Sheth (1967) offer classification schemes for theories as a device for assessing the status of marketing theory. Robin (1970) questions the ethical desirability of the developing positive science of marketing, although his ethical concerns seem limited as he focuses exclusively on the question of individual privacy as it clashes with the informational needs of a science. Dawson (1971) suggests that marketing may lose its inherent relevance as it becomes more scientific, but this paper is perhaps more valuable for its review of the concepts of Kuhn (1962). The single most comprehensive work for students of marketing is that by Zaltman (1973) and his discussion of metatheory.

A solid theory will provide several advantages. As a theory becomes well accepted, and therefore a paradigm in Kuhn's terminology, it provides the community of scholars with standard definitions for key concepts and variables, a common language of communication. Such definitions, and associated classifications, will permit an understanding of the dimensions along which results may be generalized, a need articulated by Kollat, et.al. (1970, p. 330).

Theory also permits the comprehension of large data fields and the integration of apparently diverse packages of information. For consumer behavior, like psychology, however, it is likely that "most useful theories will, for some time to come, be small conceptual systems dealing with a restricted range of phenomena", (Marquis, 1948, p. 434). The domain of behaviors that deserve explanation is simply so large that it would be optimistic and premature to expect a very powerful general theory. But hopefully theoretical developments, aided by increasing knowledge, will permit the gradual integration of such small conceptual systems into increasingly larger systems. This we do now by modelling, but the same result might be obtained in the future by the introduction of new concepts at a different level of abstraction that serve to unite and show communalities between currently distinct topic areas. As subsystems are amalgamated into a theory of consumer behavior, that theory will be increasingly integratable into an even more general theory of marketing.

Good theories are of primary value in the role they play in structuring research activities. Theories gain their acceptance by successfully "explaining" those problems perceived as most acute by the practitioners in the field--the previous anomalies. The theories also provide good definitions of problems and, importantly, offer a conceptual scheme that permits approach to those problems. Theories will be perceived as worth-while only so far as the practitioners perceive that the problems raised are solvable by operations within the theoretical structure, even if those problems are not yet solved. Without being able to presume the solvability of such problems the theory has little pragmatic value for the scientists.

Theory also facilitates specific research design, by identifying variables for control. The operation of theories generates testable hypotheses. Theories also permit the identification of needed research by, for example, identifying critical assumptions. "The more interesting and directly productive functions of theory in basic research, however, are those of coordination and planning" (Krause, 1971, p. 219). "A unified theory of consumer behavior may help to

coordinate and control the specialized units by reducing the differences in goals and perceptions ..." (Sheth, 1967, p. B718). The heuristic role that theory plays in identifying research goals, and the establishment of goal priorities is a valuable one as it permits the planning of research programs.

An Approach to the Planning of Research

To permit a science to evolve on an ad hoc basis, reflecting the idiosyncratic interests of its practitioners, the availability of data, and the convenience of measurement instruments, implicitly assumes that anarchical systems progress faster than do organized systems. Since few would hold that to be true in the political arena it is surprising that they appear to do so in the realm of science. But it is being increasingly recognized that "The organization of science is itself a scientific task, it is even a new science that has developed in our time; it is concerned with science and its laws, its development, its particularities, the general and specific aspects of its various disciplines, and its dependency on historical evolution and economic factors. In a word, science is the subject for scientific research" (Science and Synthesis, 1971, p. 151).

In the gathering of research information for the confirmation of theories one can conceive of three levels of planning and organizing activity. At the most microscopic level would be the experimental design; the design of a single experiment, although perhaps involving a number of treatment conditions of some complexity, aimed at determining the validity of one, or at most a few, assertions or hypotheses. At the other extreme, is what one might call policy design; the large scale planning of a "science", the selection of major alternative thrusts of a discipline. This is the level of planning that is at least implicit in the funding decision of large funds suppliers like the National Science Foundation. The decisions of who and what they support affect the character of the science that evolves.

But in consumer behavior there are few if any who have the opportunity and power to make decisions at that level, most of us face the more immediate research planning activities of our own private research ventures, and we have come to do so with some sophistication, at least at the first level described. Primarily because of our close association with market research techniques, we are trained to design experiments with great efficiency and precision, using the most elegant of designs, sampling techniques, measurement instruments, and the like. But there remains a level of research planning, obviously between the two described extremes, in which we do very little. That is the area of program design.

Program design is the "planning of an integrated set of projects focussed on a central problem" (Marquis, 1948, p. 431). That "problem" may for example be an hypothesis with the program aimed at discovering all of the conditions that are necessary and sufficient for the hypothesis to be valid. But it is a series of experiments, or research activities whose central thrust is some specific research problem. A fully developed research program would probably involve people with a number of disciplinary backgrounds, or at the least the problem would be attacked using a number of methodological approaches. A research program, unlike an experiment, expands the coverage of a research activity far enough to permit the examination for progress.⁵ Program research is certainly more than the bleeding of a single project or data bank for a number of publications. It is also more than repeated application of a tool in search of problems. It is also more than the exercise of trivial variations on a successful theme. Programmed research is the attempt to be purposeful in some rational

way; to address oneself to a research problem in an efficient and potentially fruitful manner. It involves replications and sequential experimental strategies (Cox, 1958). It involves all of the major phases of research activity, from problem formulation, through literature reviews, pilot testing, theoretical elaborations, to hypothesis testing. If successful, a research program generates the more information, with less redundancy of effort, time lost, operational vagueness, incomparability of data and general confusion. A well thought out research program certainly has more potential for information productivity than does an aggregation of individuals each leaping from whim to whim.

A Topological Model of Research Information

A most intriguing and productive way of thinking about the planning of research activities employs a topological conception of causal patterns, whether those patterns are supposed (theoretical) or actual (empirical). It is the kind of conceptualization that I personally find comfortable and convenient, but it also seems to be quite generally tractable and comprehensible, thereby allowing for further development and explication following this introduction to the ideas.

It seems valuable to follow the lead of Cox (1958) and to conceptualize a "response surface" which portrays the relationship between the phenomenon under scrutiny, the dependent variable, and all its causal factors, the independent variables. Such a surface is a k -dimensional generalization of a regression line, (or multiple regression plane) and the surface is simply that defined by the functional relationship between the dependent variable and all of the independent variables. Such a surface can and probably will take a variety of complex topological forms.

Let us however consider a response lattice rather than a response surface. This change hardly affects our thinking, but it does more accurately reflect the fact that many of our variables are discrete, especially independent variables that are experimentally fixed at specified levels. While the continuous model has certain appeals, it is rare that we can afford the thoroughness of research to approximate continuous measures on all of the variables. Let us also, just for the convenience of conceptual visualization, consider only three dimensional lattice with two independent variables as the base axis and the vertical elevation being the value of the dependent variable. In applying this model to an actual experimental program, however, it should be noted that it is essential that the early conceptualization of a lattice be expansive in its inclusion of factors or independent variables. This is so because the addition of a dimension to a response lattice at some date subsequent to the collection of research data makes the already collected data indeterminate along the new dimensions.

It is now possible to describe an experimental program by its location(s) within the lattice, the density and completeness of its coverage in some area, and the number of replicates undertaken. This is done by Krause (1967) who pursues the discussion to elaborate how such attributes will vary with different experimental purposes. The purposes he discusses are: the testing and qualifying or restricting a specific causal proposition ("scope restrictive"); developing a comprehensive causal proposition ("variance exhaustive"); and describing the efficacy of a specific set of treatments ("mapping"). That discussion will not be reiterated upon here. We will instead exercise this conception as a framework for discussion of research activities.

In any research program, the results of early experiments will have an influence on subsequent experiments. In fact, the conditionality of the latter experiments may be seen as a necessary component of truly programmatic research. All too many experiments are designed as if they were single-shot scientific excursions which will exhaust academic interest in an area. Recognizing that experimental exploration of a phenomenon is not so easily accomplished, suggests the importance of interpreting the early research in a programmatic manner, i.e., looking for its implications for future research. Informative early research is often informative because of what it tells us that we don't know, rather than what it tells us that we do know. The programmatic implications of early research are probably more important than the substantive findings and our evaluation of such research ought to reflect this.

In terms of the response lattice, what early experiments can do is to point the way to areas where the surface is irregular, complex, or ill-behaved. These can be thought of as "rough" areas. Early experiments also point to areas which are indeterminate, that is, where the variance of the dependent variable is high relative to the surface variation. These can be called "soft" regions. Possibly, but typically unlikely, early experiments will define surface sections that are both firm and smooth (well behaved). In these rare instances, the experiment does provide substantive information as well as programmatic.

Once the process of mapping a response surface has begun, there are a number of strategies that can be followed. Perhaps the most difficult, expensive, and maybe the least valuable overall is the exploration of the rough surfaces. This process requires fine variations of the independent variables, but only over a limited range, and precise measure of the dependent variables. Alternatively, of course, one could "explore" the soft areas. This exploration usually involves the improvement of measurement techniques to increase reliability, or the introduction of new factors. Unlike the exploration of rough regions, it may not involve extensive experimentation over small ranges. It may involve replications of experiments at existing coordinates in the lattice, varied perhaps over some new dimension that will help "explain" the observed variation in the dependent variable. The last simple strategy is just to explore the completed undefined, currently unexplored areas.

Once the form and contours of the response surface start to take shape, and once that shape is firmed up by improved measurements and the aggregation of sample information, several strategic alternatives present themselves. Research programs might try to find maxima or minima, as is commonly done. They might also go through a search for sections of the surface that are the steepest, especially the steepest around a maxima. Such research answers the questions of "along what dimensions, or combination of dimensions, does the value of the dependent variable change the fastest?" or "to what independent variables is the phenomenon most sensitive?" One might also look for isometrics, the contour lines of the map of the surface to find out what alternative combinations of independent variables generate the same value of the dependent variable. Such a tactic, like the maxima search, might be valuable in pragmatic research as it would ultimately permit the more effective allocation of resources by creating what are in some sense indifference curves. Lastly, although this probably does not exhaust the alternatives, one might search for "firm" islands within soft regions. This identifies those sections of the surface where one can more comfortably predict what the relationship is between the independent and dependent variables.

The ability to pursue a program of research for the mapping of some response surface depends on measurement capabilities and the cost of research.

It also depends, in a more basic way, on theory. Theory provides the research programs with suggestions about the important variables for inclusion in the factor lattice, suggestions about the locations of maxima, steepest ascents, and other topological features. It is rare, however, for a theory to be an explicit statement about the nature of the response surface. Many theoretical statements are certibus paribus in nature, describing the intersection of the response surface and a plane parallel to one of the axes. Or they may be generalizations, a description of the projection of the surface on one of the coordinate "walls". But even though the theories are rarely explicit statements of the functional relationship describing the surface, theories do provide direction for search. They do so by providing the researcher with an a priori mapping which is first a guide and then a foundation for the revision of that map in light of subsequent research results.

Theories, unlike maps however, always suggest features of yet uncharted regions. If they do not do so they are not valuable as theories and are only language transformations of existing knowledge, and not transformations into a more powerful language. The combination of good theory with a model for research planning is a powerful one, and it can both permit and encourage greater efficiency of research efforts.

Footnotes

1. Richard W. Pollay is Associate Professor of Commerce and Business Administration at the University of British Columbia, and the Editor of the Journal of Business Administration.
2. This is, unfortunately, the case in most academic areas these days and is as much a reflection of the publish or perish criteria used by university promotion and tenure committees as it is a reflection on the state of the art in consumer research.
3. The text provides no argument that any professional insecurity exists within the consumer behavior field, although impressionistic observation suggests that it certainly does exist. An argument, admittedly weak in both structure and evidence, can be made. The tendency of the university professional in this area to publish for the sake of publishing, to be a professional academic, and to use publishability as his criterion of scientific value may be seen as indicative of professional insecurity as it betrays a lack of criteria intrinsic to the science. This insecurity may often be masked, however, by the tendency to pass flippant judgements on the value of others research, judgement often based on either the esotericism or the familiarity of the alternative research and not on conclusiveness precision or value of its results.
4. Dawson (1971) erroneously concludes that marketing is a crisis science. In doing so he failed to recognize that to qualify as a crisis science there needs to have been a normal science with a well accepted paradigm, and the challenge of this paradigm by a competitive one. Since neither marketing in general, nor consumer behavior in specific, have either the history of a paradigm or a well articulated competitor, Dawson's categorization seems inappropriate. Dawson's confusion is also seen in his expression of fear that we "will lapse into the practice of normal science", (p. 72). (Emphasis added.)

5. This is an explicit criterion because of the notorious failure of the academic community to be motivated by statements such as "what is now required is further research". Indeterminate research, which is the kind most frequently employing such a call for more research, is of no appreciable value and its unimpressive character is hardly a source of inspiration. The failure of other researchers to take up the lead provided is probably also the result of a once common practice of academicians to stake out a claim on a research problem by asserting that subsequent research was underway, even when the assertions had no basis in fact.

References

- Cox, D.R. The Planning of Experiments. New York: Wiley, 1958.
- Dawson, L.M. Marketing Science in the Age of Aquarius. Journal of Marketing, 1971, 35, pp. 66-72.
- Engel, J.F., Kollat, D.R. & Blackwell, R.D. Consumer Behavior. New York: Holt, Rinehart and Winston, 1968.
- Halbert, Michael H. The Requirements for Theory in Marketing. In Cox, R. et al., (eds.) Theory in Marketing. Homewood, Illinois: Richard D. Irwin, 1964, pp. 17-36.
- Howard, J. & Sheth, J. The Theory of Buyer Behavior. New York: John Wiley, 1969.
- Hunt, S.D. The Morphology of Theory and the General Theory of Marketing. Journal of Marketing, 1971, 35, pp. 65-68.
- Kollatt, B.T., Engel, J.F. & Blackwell, R.D. Current Problems in Consumer Behavior Research. Journal of Marketing Research, 1970, 1, pp. 327-332.
- Krause, Merton S. Proving Causal Propositions: The Foundations of Program and Experiment Design. Multivariate Behavioral Research, July, 1967, pp. 349-376.
- Krause, Merton S. Corroborative Results and Subsequent Research Commitments. Journal of General Psychology, 1971, 84, pp. 219-227.
- Kuhn, T. S. The Structure of Scientific Revolutions. Chicago: University of Chicago Press, 1962.
- Lee, C. E. Measurement and the Development of Science in Marketing. Journal of Marketing Research, 1965, 2, pp. 40-ff.
- Marquis, D. G. Research Planning at the Frontiers of Science. American Psychologist, 1948, 3, pp. 430-438.
- Nicosia, F. M. Consumer Decision Processes, Englewood Cliffs, N.J.: Prentice-Hall, 1966.
- Perloff, Robert. Consumer Analysis. Annual Review of Psychology, 1968, 19, pp. 437-466.
- Polanyi, Michael, Personal Knowledge: Towards a Post-Critical Philosophy. Chicago: University of Chicago Press, 1958.
- Robin, D. P. Toward a Normative Science in Marketing. Journal of Marketing, 1970, 34, pp. 73-76.
- Science and Synthesis. New York: Springer-Verlag, 1971. See Part 2, Ch. 4, The Organization of Scientific Research, pp. 147-178.
- Sheth, Jagdish N. A Review of Buyer Behavior. Management Science, 1967, 13, pp. B718-B756.
- Zaltman, Gerald. A Metatheory of Consumer Behavior, Forthcoming, (1973).

SOCIAL CLASS AND INCOME INFLUENCES ON EXTERNAL
SEARCH PROCESSES OF ADOLESCENTS

Stephen K. Keiser
University of Delaware
and Philip G. Kuehl¹
University of Maryland

Introduction

In recent years, the adolescent segment of the population of the United States has received increased attention and study from many societal institutions. For example, political parties, religious groups, educational institutions, and governmental agencies have attempted to study the impact of adolescent behavior and attitudinal structures on the performance of their tasks and activities. In a similar manner, producers of economic goods and services have recognized the importance of understanding the "adolescent market" for the following four reasons:

1. Adolescents comprise an important, unique market segment for many types of goods and services with an estimated spending level in excess of \$20 billion annually (National Industrial Conference Board, 1969).
2. Adolescents are estimated to affect the expenditure of \$60 billion annually through their influence on the consumption patterns of other family members (Senior Scholastic, 1965).
3. The market behavior of adolescent groups, such as shopping behavior and brand preference, affects their consumption behavior as adults.
4. Adolescent behavior has affected the consumption behavior of all members of the American society. For example, their influence has been observed in clothing and hair styles in recent years.

One important dimension of the consumption behavior of adolescents which is of interest to researchers and practitioners are the processes by which adolescents acquire and utilize information about their environment. Unfortunately, previous research on the information search behavior of adolescents is limited in the following manner:

1. Research has been limited to the investigation of only one or two information sources (Ward & Robertson, 1970; Bowerman & Kinch, 1959; and Editor and Publisher, 1966). As a result the information obtained from these studies is of little use in explaining the relationships between the use of all possible sources of information.
2. Many of the previous studies have concerned aspects of adolescent behavior other than consumption behavior (Remmers & Radler, 1957).
3. Past research has treated adolescents as a homogeneous group. As a result the differences in the consumption behavior of adolescents with different economic resources or social status have been ignored (Gilkinson, 1965).

The general objective of this study, then, is to examine the processes by which information about economic goods and services is obtained by adolescent consumers. These processes are referred to as external search behavior. The nature of these processes are examined from the viewpoint of (a) the relations between different information sources and (b) the effect of the socio-economic characteristics of adolescents on search behavior. Two characteristics of con-

sumers that have been shown to influence external search processes are social class and income. The strong research tradition of these two predictor variables suggests that they should be effective predictors of external search behavior of adolescents.

Research Objectives

The objective of this study is to investigate the extent to which two predictor variables, income and social class, are related to the external search processes of the adolescent segment of American society. In specific terms, the research focused on:

1. Examining whether or not external search processes can be understood in terms of differences between adolescents with different incomes.
2. Examining whether or not external search processes can be understood in terms of differences between adolescents in different social class categories.

As indicated above, income and social class are the two predictor variables investigated in this study. Using the Engel, Kollat, and Blackwell conceptual model as a framework for the research, the following sources of information served as the major criterion variables in this study:

1. Mass media sources: including the degree of utilization of broadcast and print media sources; and, the relationship of brand awareness to external search processes.
2. Personal sources: including the examination of opinion leadership, family, and reference groups as sources of information.
3. Marketer-dominated sources: including the impact which retail store shopping behavior has on external search activity.

The degree of utilization and importance of these information sources for adolescents constitutes the major empirical thrust of this study. In order to accomplish these objectives, a questionnaire was designed, pretested, and administered to a sample of 1,200 junior high and senior high students in Columbus, Ohio in March, 1971.

Limitations of the Study

The use of Columbus, Ohio students as the sampling frame represents two obvious limitations of the study in terms of generalizing these results to adolescents in other geographic areas. First, the researchers did not attempt to include (a) adolescents not enrolled in Columbus, Ohio secondary schools, or (b) students absent on the day the questionnaire was administered to the sample. In this respect, the sample tends to be representative only of students enrolled and attending class on the day the questionnaire was administered. However, the questionnaire was administered on a day when absenteeism tended to be lowest (Thursday) and State of Ohio law requires that adolescents must be enrolled in school until age 17. These considerations tend to minimize the seriousness of this first limitation.

The second limitation in the study evolves from the fact Columbus, Ohio is not representative of the national population. However, age, income, educational, and occupational characteristics of Columbus, Ohio are similar to those of the United States as a whole. As a result, a carefully constructed sample of Columbus, Ohio adolescents does tend to be somewhat representative of adolescents in general.

Another limitation of this study evolves from the nature of the research design used in the research. Inasmuch as the research design used in the present study was cross-sectional, rather than longitudinal, the results of the study should be used cautiously when predicting changes in behavior over-time.

Finally, the results cannot be generalized to specific products because length constraints on the questionnaire eliminated the opportunity to ask adolescents about their external search processes for specific products. As a result, respondents were asked about "things they buy" rather than specific products.

Conceptual Framework

This section discusses the conceptual framework and variables used in the study. The Engel, *et al.* consumer decision-making model, which was adopted as the conceptual framework, is discussed first. External search behavior is then defined and discussed. An explanation of the dimensions of external search behavior examined in the study are included in this discussion. Next, income and social class, the predictor variables, are operationally defined. This is followed by an explanation of the means by which the different research variables were measured.

Consumer Decision Making as a System

As the consumer behavior "research tradition" develops and gains in methodological sophistication, the need for conceptual models to guide such research becomes apparent. In order to incorporate the results of the present study within the context of the existing consumer behavior "research tradition", the Engel *et al.* consumer decision-making model was adopted as the conceptual framework.

Their model, as shown in Figure I, describes the consumer decision process as a system with outputs that result from the processing of received inputs. This cognitive processing of environmental inputs is directed by the component labelled the "central control unit" which is the memory and thinking of the consumer. The output of consumer cognitive processes is visualized as consisting of five "decision steps": (1) problem recognition; (2) external search for alternatives; (3) evaluation of alternatives; (4) purchasing processes; and (5) postpurchase behavior. In the present study, the external search for alternatives stage of the decision process is examined in relation to the effects of the adolescents income and social class on the central control unit.

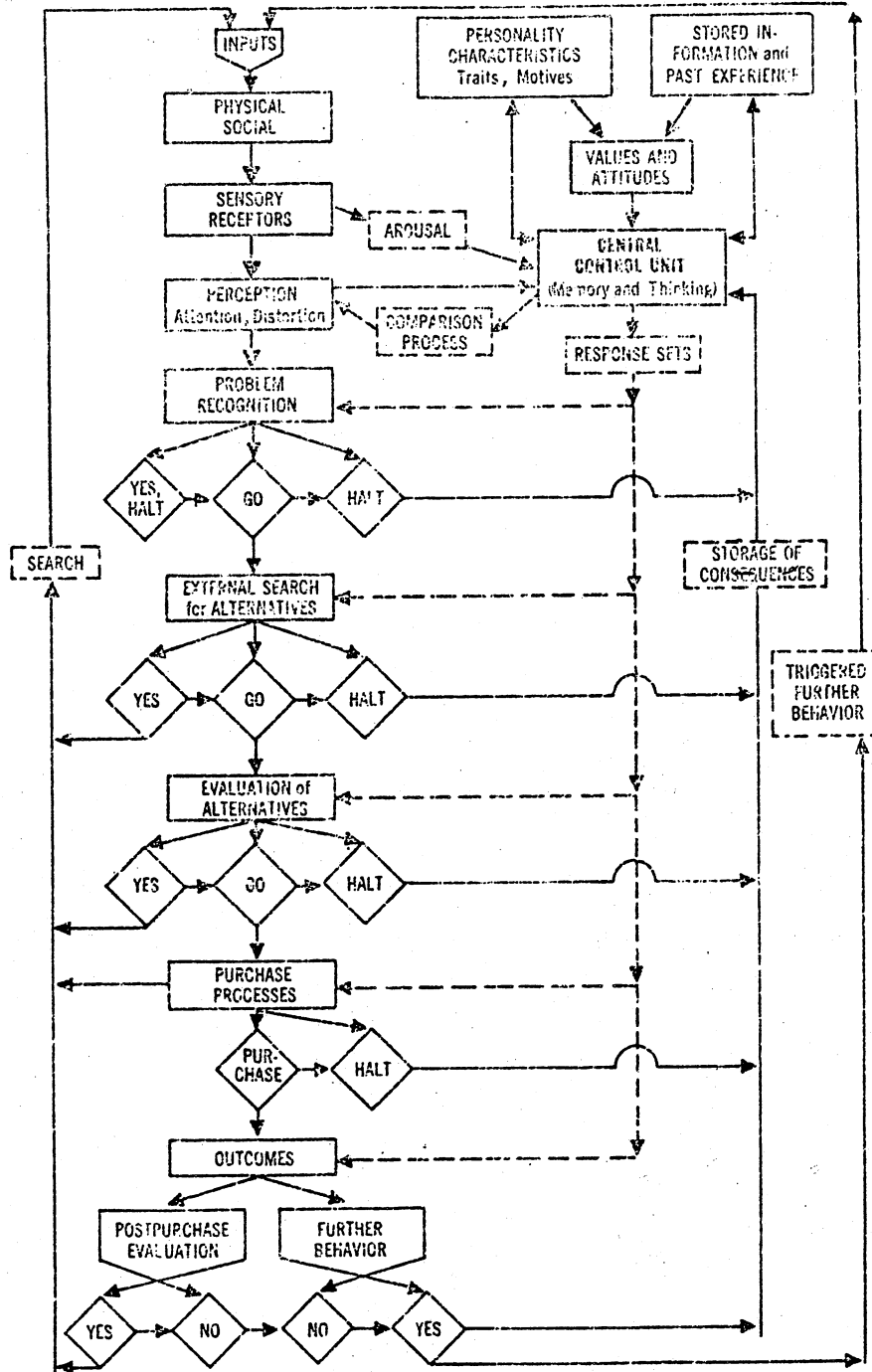
External Search Behavior--Defined

External search behavior is defined in this study as:

...processes and activities whereby the consumer uses various sources of information, including mass media, personal sources, and marketer-dominated sources...to learn about the number of alternative solutions to the perceived problem, the characteristics and attributes of those alternatives, and their relative desirability (Engel, Kollat, & Blackwell, 1968, p. 378).

In specific terms, the primary focus of this study is the examination of the following sources of information used by adolescents in external research.

Figure I--A COMPLETE MODEL SHOWING OUTCOMES OF THE PURCHASING DECISION



Source: James F. Engel, David T. Kollat, and Roger D. Blackwell, Consumer Behavior (New York: Holt, Rinehart, and Winston, Inc., 1968), p. 50.

1. Mass media sources: print and broadcast media which reach large number of consumers.
2. Personal sources: members of the extended family, friends, school teachers, and strangers who contribute information to any given consumer.
3. Marketer-dominated sources: the marketing activities of various firms by which they attempt to communicate with consumers, such as sales presentations by shopkeepers or clerks, in-store displays, consumer visits to retail outlets, and catalogs mailed to prospective buyers.

Figure II, on the following page, presents a detailed summary of these information sources.

Dimensions of External Search Behavior Investigated

The use of information sources in external search can take many forms. These information sources may (a) initiate the decision-making process; (b) influence the exact nature of the decision-making process such as time of purchase; or, (c) place of purchase. Several of these dimensions of external search behavior are investigated in the present study.

All Information Sources

Some of the dimensions of external search behavior investigated involved all of the categories of information sources described above. These included the use of information sources to ---

- a. initiate decision-making processes related to consumption. This aspect involves the sources which are the first to make the consumer aware of the product. The relative frequency with which the different information sources were used for initial consumption information was investigated.
- b. to provide information subsequent to the initiation of the decision-making process. Similar to initial product information, respondents' relative use of the information sources for subsequent product information was examined.
- c. provide decision rules used to solve the various aspects of the decision-making process. This role of information sources is one of "decisive effectiveness" (Engel, Kollat, & Blackwell, 1968, p. 404). The present study investigates the information sources individuals consider the most influential and the least influential.

Mass Media Sources

One aspect of external search processes investigated here relates to mass media. This aspect was measured by the amount of exposure to mass media. Mass media exposure is divided into (a) broadcast media exposure and (b) print media exposure. Broadcast media exposure is measured as the number of hours spent by a respondent on an average day watching television or listening to radio. In anticipation of variations in broadcast exposure patterns for different times during the week, a respondent's exposure to television and radio was investigated for two time periods--weekdays and weekends. Print media exposure was defined for purposes of the study as a combination of the (a) number of newspapers and magazines read by a respondent and the (b) frequency with which he

read the media. Mass media exposure provides one indicant of the use of information sources in external search.

FIGURE II

A DELINEATION OF INFORMATION SOURCES USED FOR
EXTERNAL PRODUCT SEARCH BEHAVIOR

- I. MASS MEDIA SOURCES
 - A. Broadcast Media
 - 1. Television
 - 2. Radio
 - B. Print Media
 - 1. Newspapers
 - 2. Magazines
- II. PERSONAL SOURCES
 - A. Members of the Extended Family
 - 1. Father
 - 2. Mother
 - 3. Brothers
 - 4. Sisters
 - 5. Relatives outside the household
 - B. Friends
 - C. Teachers
 - D. Strangers
- III. MARKETER-DOMINATED SOURCES
 - A. Store-Related
 - 1. Shopkeepers or clerks
 - 2. Store displays
 - 3. Store visits
 - B. Mail Catalogs

Personal Sources

Many types of personal information sources are available to adolescent consumers during the decision process. For example, one important source of information influencing the consumption of adolescents is friends. In view of the importance of friends on adolescent consumption behavior, the role of friends as opinion leaders was investigated. Opinion leaders are defined as "those to whom others turn for information and advice" (Robertson, 1970, p. 83).

Marketer-dominated Sources

Two dimensions of the role of marketer-dominated sources in external search were investigated. One of these dimensions investigated the overall effectiveness of messages sent by marketers through other information sources. For example, external search behavior may result in the learning of brands and advertising slogans. As a result, "brand awareness" was adopted as another indicator of the nature of external search behavior. For purposes of this study, "brand awareness" is defined as: (1) the ability to correctly associate brand names and producers, and (2) correctly complete or identify the sponsor of popular advertising slogans.

The other aspect of marketer-dominated information sources investigated involved retail store patronage. It is especially pertinent for the retailer

to understand how adolescents choose one store rather than another as a means to seek information about products. As a result of the importance of store visitation patterns of adolescents to the retailer, the relative importance of the following characteristics in the store choice decision were investigated.

1. Convenience of the store.
2. Trustworthiness of store personnel.
3. Friendliness of store personnel.
4. Quality of store offering.
5. Availability of latest styles.
6. Service provided by the store.
7. Assortment available.
8. Low prices.

External Search--Limitations

The processes of external search behavior are examined in this study within the following constraints.

1. No assumptions were made regarding whether the respondents in the study are involved in:
 - a. extensive problem solving,
 - b. limited problem solving, or
 - c. routinized response behavior (Howard & Sheth, 1967).
2. No assumptions were made regarding the source of the consumers' problems, such as assortment deficiency, expanded consumer desires, or expanded consumer demands (Walters & Paul, 1970).
3. No assumptions were made concerning the
 - a. perceived value of the external search,
 - b. the appropriateness of stored information,
 - c. the ability to recall stored information,
 - d. the degree of perceived financial, social, or physical risks in purchasing the products (Engel, Kollat, & Blackwell, 1968, p. 382-385).
4. No assumptions were made concerning the real or perceived cost of the external search (Engel, Kollat, & Blackwell, 1968, p. 382).
5. The sources of information used by consumers in the external search process are researched within the context of the Engel et al. (1968, p. 34-53).

Income as a Variable

One of the first variables used by early researchers to explain consumer behavior was the buyer's income. In an effort to measure the adolescent's spending power regardless of its source, income of adolescents is defined as: the amount of money the adolescent perceives himself receiving during an average week from all possible sources of earnings including allowances and occupation pursuits.

Social Class as a Variable

The use of social class as a market variable was first suggested by Martineau in 1958, and it has been treated as a major market variable by many writers for the past decade. Early researchers in the use of social class claimed or implied that social class was a better variable than income to use when explaining buyer behavior. For example, Coleman wrote, concerning furni-

ture purchases of consumers in different social classes, that: "The correlation between prices of goods purchased and social class is relatively quite high in these product areas while the correlation between price paid and annual income is lower than one might expect (Coleman, 1960). Subsequent research has cast doubt on the assumption that social class is a better predictor of consumer behavior than income (Slocum & Mathews, 1970; Rich & Jain, 1968; and Myers, Stanton, & Haug, 1971). The reason social class may be losing its predictive power is that: "...social class distinctions have been obscured by rising income and educational levels" (Rich & Jain, 1968). Since previous studies suggest that income and social class have different effects on consumer decision making, the two variables are treated as separate independent variables in this study.

Social classes are defined for this study as: "...relatively permanent and homogeneous divisions in society into which individuals or families can be categorized when being compared with other individuals or families in the society (Engel, Kollat, & Blackwell, 1968, p. 264).

Conceptual Framework--Summary

In summary, the study was conducted within the Engel, et al. decision process framework and involves an investigation of the relationships between selected predictor and criterion variables. The criterion variables are:

1. income.
2. social class.

The predictor variables, which encompass different aspects of external search processes, are:

1. the use of different information sources to initiate decision-making processes related to consumption.
2. the use of different information sources to provide information subsequent to the initiation of the decision-making process.
3. the types of information sources considered to be the most and the least influential in consumption-related decisions.
4. amount of exposure to mass media.
5. opinion leadership.
6. brand awareness.
7. retail store patronage motives.

Measurement of Variables

The types of questions framed to measure each of the variables are shown in Figure III on the following page. A mixture of open-ended and multiple-choice questions were employed. The exact nature of these questions is summarized in the third column of Figure III.

Methodology

Data for the sample were collected in March, 1971 through a questionnaire form completed by a sample of 1,280 junior high and senior high school students in Columbus, Ohio. The questionnaire form was pretested in York County, Pennsylvania and Prince Georges County, Maryland, before its use in Columbus, Ohio. It was administered to the respondents with the cooperation and assistance of local school guidance counselors and administrators. The questionnaire form took approximately 45 minutes (one class period) to complete.

FIGURE III
SUMMARY OF MEASURES

<u>Variable</u>	<u>Format of Question</u>	<u>Nature of Question</u>
Income	Open-ended	Amount of earnings in an average week
Social Class (Hollingshead's Index of social position)	Open-ended (Father's Occupation) Multiple-Choice (Level of Father's education)	Father's occupation Educational level
Initial information	Multiple-Choice	5-point "always - never" items, such as: "How often is your father the first source to provide you with information about things you buy?"
Subsequent information	Multiple-Choice	5-point "always - never" items
Decisive information	Multiple-Choice	Which source do you consider most (least) important?
Broadcast media exposure	Multiple-Choice	Amount of time spent watching television (listening to radio) on weekdays (on weekends)
Print media exposure	Multiple-Choice	6-point frequency of readership items
Opinion leadership	Multiple-Choice	3-point items concerning likelihood of being asked by others for information about products
Brand Identification	Open-ended	20 brand names such as: Vega, and respondent was asked to provide name of producer
Slogan recall	Open-ended	10 incomplete slogans such as: Fire yours, Hire ours
Retail Store patronage motives	Multiple-Choice	Rank 8 motives in relative importance

The respondents were selected through the use of a multi-stage sampling process using individual schools and grade sections as the primary sampling units. In this sampling process, all secondary schools in the Columbus, Ohio system were divided into junior high and senior high school stratifications. Next, five schools were randomly selected from each strata. Then each of the 10 schools included in the sample were stratified by grade level and 2 home room sections of students were randomly selected from each grade strata. The questionnaire was administered to all students present in the selected home rooms during the home room period at each school.

The data provided by the 1,280 respondents were used to accomplish the objectives of the study. In order to determine if the differences observed in the external search processes of adolescents were attributable to sampling error or to actual differences in behavior, several statistical tests were adopted. The three statistical tests used in this study were the (a) Chi-square test, (b) Wilcoxon matched-pairs signed-ranks test, and (c) Friedman two-way analysis of variance. All three tests are nonparametric tests. This type of statistical test was deemed more appropriate than parametric tests because the data were assumed to be abnormally distributed and ordinal in nature.

The Chi-square test was used to examine the relationships existing between the predictor variables and most of the criterion variables. The relative rank of retail patronage motives was the only criterion variable which was not examined with the Chi-square test. The relationships existing for this criterion variable were tested with the use of the Friedman statistic. The Wilcoxon test was used to determine if the relative variations in use of different information sources were significant.

The significance of the observed relationships for each of the test was determined by either the value of the Chi-square distribution or values of the Z distribution. The Chi-square distribution was used to determine the significance of the Chi-square and two-way analysis of variance statistics. The Z distribution was used in the case of the Wilcoxon test.

An important issue involved in statistical analysis is the definition of the "region of rejection". The region of rejection is a subset of the sampling distribution that the researcher feels is so extreme that when the research hypothesis is true the probability is very small that the sample actually observed will yield a value which falls into this region. Inasmuch as there is little theory to guide the researcher in establishing the size of the rejection region the level established by research tradition, 5 percent, was adopted for the statistical analyses of this data.

Research Results

As indicated in the preceding discussion, three general types of information sources (mass media, personal, and marketer-dominated) were investigated in this study. The results are organized using these three types of information sources. For each of these types of information sources, the dimensions of external search processes (initial, subsequent, and decisive impact) are discussed as well as their relationships with income and social class.

Mass Media

The first class of information sources investigated were mass media sources. As previously discussed, television, radio, newspapers, and magazines were the four specific sources included in this investigation.

Table I
Relative Frequency of Use of Information Sources
For Initial Product Information

Source	Mean Frequency of Use	Rank				Marketer-Dominated Sources
		All Sources	Personal Sources	Mass Media Sources		
Mother	3.27	1	1			
Store Visits	2.86	2			1	
Friends	2.78	3	2			
Television	2.58	4		1		
Newspapers	2.46	5		2		
Father	2.45	6	3			
Radio	2.44	7		3		
Magazines	2.41	8		4		
Store Displays	2.38	9			2	
Mail Catalogs	2.24	10			3	
Sisters	2.14	11	4			
Brothers	1.91	12	5			
Shopkeepers or clerks	1.82	13				4
Relatives outside the household	1.76	14	6			
Teachers	1.47	15	7			
Strangers	1.26	16	8			

Table II

Results For Hypothesis H₁ Comparing Relative Utilization of
Information Sources For Initial Product Information

Information Sources Tested	Computed Wilcoxon T Value	Computed Z Value	Level of Significance
Mother - Store Visits	89,463.5	6.87	.001
Mother - Friends	80,932.0	8.89	.001
Store Visits - Friends	91,306.0	1.96	.030
Store Visits - Television	73,065.0	5.87	.001
Friends - Television	80,228.5	4.48	.001
Friends - Newspapers	74,924.5	6.76	.001
Television - Newspapers	33,813.0	3.65	.001
Television - Father	88,198.0	2.95	.002
Newspapers - Father	92,777.0	.51	.301
Newspapers - Radio	41,830.5	.54	.290
Newspapers - Magazines	36,524.0	1.71	.050
Father - Store Displays	102,625.0	.92	.210
Father - Radio	103,977.0	.11	.460
Father - Magazines	98,279.5	.45	.330
Father - Mail Catalogs	89,844.5	3.74	.001
Radio - Magazines	49,375.0	1.03	.150
Radio - Store Displays	65,791.5	1.15	.130
Radio - Mail Catalogs	65,396.5	4.61	.001
Magazines - Store Displays	60,346.5	.63	.260
Magazines - Mail Catalogs	54,490.0	4.18	.001
Store Displays - Mail Catalogs	50,540.0	3.94	.001
Store Displays - Sisters	67,248.5	12.44	.001
Mail Catalogs - Sisters	83,413.5	1.65	.005
Mail Catalogs - Brothers	62,256.0	6.51	.001
Sisters - Brothers	43,466.5	4.89	.001
Sisters - Shopkeepers or clerks	60,540.0	6.32	.001
Brothers - Shopkeepers or clerks	68,450.0	1.82	.030
Brothers - Relatives out- side the house- hold	56,605.0	3.38	.001
Shopkeepers or clerks - Relatives outside the household	67,070.5	1.71	.040
Shopkeepers or clerks - Teachers	24,424.0	9.58	.001
Relatives outside the household - Teachers	26,279.5	8.54	.001
Relatives outside the household - Strangers	13,481.0	14.28	.001
Teachers - Strangers	9,939.5	8.23	.001

Initial Product Information

The average frequency with which respondents used the different information sources for initial product information were used to rank the different types of sources as shown in Table I. The average for each source was determined by scaling responses on a one (never) to five (always) scale and summing these values for all respondents. Then, this sum was divided by the size of the sample.

As indicated in Table I, television is the mass media source used most frequently for initial product information while magazines were used least frequently. In order to determine if the differences in frequency of use of different information sources were significant, two statistics, the Wilcoxon T and Z, were computed for pairs of information sources. These information sources were paired on the basis of their relative rankings reported in Table I. The values of the T and Z statistics are reported in Table II. It is concluded that television is used significantly more frequently for initial product information than other mass media sources with the Z value computed for the pair of mass media sources, television and newspaper, significant at a .001 alpha level (See Table II). Conversely, magazines were not used significantly less frequently for initial product information than was radio although the difference between magazines and the other mass media source, newspapers, was significant.

Income. The Chi-square values which resulted when the effect of income was related to the use of information sources for initial product information reported in Table III. In addition the results are reported for the criterion variable of social class in this table.

Income does not have a significant effect on respondents' utilization of mass media sources for product information. Subsequent analysis of the two sample strata, junior high and senior high school, revealed that the higher the income of junior high respondents the more they utilized magazines for initial product information ($\chi^2 = 14.16$, level of significance = .05). In all other cases the use of mass media for initial product information was not related to the amount of respondent income.

Social class. Mass media are not utilized differently for initial product information by adolescents who are in different social classes (See Table III). This lack of relationships was also found to be the case when junior high and senior high respondents were considered separately.

Subsequent Product Information

Similar to the analysis of initial product information behavior, average frequency of use of information sources for subsequent product information was used to rank the sources, as shown in Table IV, and these rankings were paired prior to performance of the Wilcoxon test (these results are shown in Table V).

Examination of Tables IV and V result in the conclusion that television is the mass media source utilized most frequently for subsequent as well as initial product information. Similarly, magazines were utilized less frequently than any other mass media.

Income. Income of respondents was not related to the frequency of use of mass media for subsequent product information (See Table VI). When junior high

Table III

Information Sources Used for Initial Product Information
and Income and Social Class of Respondent

Predictor Variable				
Information Source	Income		Social Class	
	χ^2	Level of Significance	χ^2	Level of Significance
<u>Mass Media</u>				
Television	4.80	.70	2.82	.70
Radio	5.22	.70	.68	.98
Newspapers	3.19	.80	2.27	.70
Magazines	10.86	.10	5.04	.30
<u>Personal</u>				
Father	16.13	.02	1.31	.90
Mother	8.48	.30	4.09	.70
Brothers	3.58	.80	1.09	.90
Sisters	4.60	.70	4.61	.50
Relatives outside the household	11.70	.10	4.58	.50
Friends	10.44	.20	13.01	.02
Teachers	5.94	.50	6.58	.20
Strangers	10.04	.20	2.39	.70
<u>Marketer-Dominated</u>				
<u>Shopkeepers or Clerks</u>				
Store displays	6.25	.50	4.62	.50
Store visits	6.17	.50	4.26	.50
Mail catalogs	9.88	.20	14.60	.01
	3.19	.80	5.06	.30

respondents were isolated, income was found to be related to the frequency of use of magazines for subsequent product information in a manner similar to that observed for use of magazines for initial product information. As junior high respondents earned higher incomes their use of magazines for subsequent product information increased ($\chi^2 = 15.73$, level of significance = .02). As a result magazine acquisition by junior high adolescents may depend on the amount of income the adolescent earns.

Social class. As shown in Table VI social class was not significantly re-

Table IV

Relative Frequency of Use of Information Sources
For Subsequent Product Information

Source	Mean Frequency of Use	Rank			
		Sources	Personal Sources	Mass Media Sources	Marketer- Dominated Sources
Mother	3.05	1	1		
Store Visits	3.03	2			1
Friends	2.84	3	2		
Television	2.73	4		1	
Newspapers	2.61	5		2	
Store Displays	2.60	6			2
Radio	2.59	7		3	
Magazines	2.56	8		4	
Mail Catalogs	2.44	9			3
Father	2.39	10	3		
Sisters	2.12	11	4		
Shopkeepers or clerks	2.06	12			4
Brothers	1.99	13	5		
Relatives outside the household	1.87	14	6		
Teachers	1.56	15	7		
Strangers	1.30	16	8		

Table V

Results for Hypothesis H₁ Comparing Relative Utilization of
Information Sources For Subsequent Product Information

Information Sources Tested	Computed Wilcoxon T Value	Computed Z Value	Level of Significance
Mother - Store Visits	186,767.0	.53	.290
Mother - Friends	140,216.0	5.09	.001
Store Visits - Friends	135,889.0	4.57	.001
Store Visits - Television	120,447.5	6.72	.001
Friends - Television	151,310.0	2.62	.004
Friends - Newspapers	126,981.0	5.59	.001
Television - Newspapers	72,665.0	3.87	.001
Television - Store Displays	96,211.5	3.87	.001
Television - Radio	72,665.5	3.87	.001
Newspapers - Store Displays	113,008.0	.28	.380
Newspapers - Radio	70,272.0	.69	.240
Newspapers - Magazines	58,641.0	1.81	.040
Store Displays - Radio	115,082.0	.33	.370
Store Displays - Magazines	97,554.0	1.46	.070
Store Displays - Mail Catalogs	82,971.0	4.76	.001
Radio - Magazines	74,516.5	1.05	.150
Radio - Mail Catalogs	111,796.0	3.85	.001
Magazines - Mail Catalogs	98,018.5	2.93	.002
Magazines - Father	138,097.5	3.87	.001
Mail Catalogs - Father	171,663.0	1.17	.120
Mail Catalogs - Sisters	18,609.5	21.70	.001
Father - Sisters	123,512.0	5.60	.001
Father - Shopkeepers or clerks	114,852.5	8.10	.001
Sisters - Shopkeepers or clerks	135,328.5	1.82	.040
Sisters - Brothers	81,920.0	2.78	.003
Shopkeepers or clerks - Brothers	132,205.0	1.56	.060
Shopkeepers or clerks - Relatives outside the household	105,711.0	5.13	.001
Brothers - Relatives outside the household	108,518.5	3.12	.001
Brothers - Teachers	61,080.5	10.75	.001
Relatives outside the household - Teachers	54,896.5	9.36	.001
Relatives outside the household - Strangers	29,928.0	17.31	.001
Teachers - Strangers	28,897.0	9.28	.001

lated to subsequent use of mass media for product information. This lack of significant relationships also held for both junior high and senior high respondents. As a result, the behavior of adolescents does not concur with that found for adults in the middle class who have been found to be more dependent on mass media for subsequent product information than members of other classes (Rainwater, Coleman & Handel, 1959).

Decisive Product Information

Television was found to be the mass media which had the most influence on external search behavior of adolescents. This finding agrees with the primary position found for television for initial and subsequent product information. Conversely the source indicated as the least important mass media source was not magazines (as expected from observation of initial and subsequent use of mass media sources). Instead, radio was considered the least important mass media source by more respondents than any other mass media source. As would be anticipated, fewer respondents considered television as the least important mass media source than any other mass media source.

Income. No significant relationships were found between income of respondents and the importance attached to mass media sources. As a result the general profile of the sample would seem to fit respondents with different incomes. This is also true for junior high and senior high respondents.

Social Class. Similarly a lack of significant relationships between social class and the influence of mass media lead to the conclusion that adolescents regardless of social class are most influenced by television and least influenced by radio. The significant differences were that senior high respondents in the upper class placed more importance on newspapers than those in other social classes. While senior high, lower class respondents were more influenced by radio than were other respondents.

Mass Media Exposure

One aspect of external search behavior examined in this study applies only to mass media information sources. This aspect is a measurement of the amount of exposure to mass media.

The correlation coefficients derived for each of the relationships for exposure to the different mass media are shown in Table VII. As noted above six types of mass media exposure were studied as both television and radio exposure were examined for two time periods--weekdays and weekends. An examination of Table VII shows that there is very little correlation between exposure to television and exposure to radio. Only when radio and television exposure for weekdays is considered does significant correlation result. Low correlation is also observed between print media exposure and broadcast media exposure. One difference is noted when the relationship between newspaper and magazine exposure is examined. Readership of these two types of print media is significantly correlated.

Income and social class. The exposure habits of respondents were related to their income and social class and analyzed using Chi-square tests. The results of these analyses show that exposure to broadcast media is dependent on the income but not the social class of the respondent (See Table VIII). The number of hours spent watching television and listening to radio on weekends decreased as the income of the respondents increased. The relationship found

Information Sources Used For Subsequent Product
Information and Income and Social Class of Respondent

Predictor Variable				
Information Sources	Income		Social Class	
	χ^2	Level of Significance	χ^2	Level of Significance
<u>Mass Media</u>				
Television	4.43	.70	.91	.95
Radio	9.31	.20	4.99	.30
Newspapers	5.42	.50	2.94	.70
Magazines	5.88	.50	3.61	.50
<u>Personal</u>				
Father	11.90	.10	7.09	.20
Mother	13.48	.05	3.41	.50
Brothers	5.81	.50	5.66	.30
Sisters	6.41	.50	2.69	.70
Relatives outside the household	3.80	.80	3.63	.50
Friends	20.86	.01	11.57	.05
Teachers	4.76	.70	1.07	.90
Strangers	4.45	.70	6.29	.20
<u>Marketer-Dominated</u>				
Shopkeepers or Clerks	12.69	.05	6.29	.20
Store displays	2.42	.90	2.51	.70
Store visits	14.30	.05	16.51	.01
Mail catalogs	2.92	.90	3.27	.70

between income and television viewing of adolescents is consistent with research examining adult behavior (King & Summers, 1971). Conversely an expected relationship of increased radio listening by respondents with higher incomes only materialized in the patterns found for weekdays.

Interestingly, the significant relationships between income and television exposure were found to be true for senior high but not junior high respondents. This may represent a shift in source of adolescent income as the adolescent becomes older. Senior high adolescents may rely more on after-school jobs for income while junior high respondents depend on income provided by

Table VII
Correlation Between Types of Mass Media Exposure

Media Exposure	Television-Weekdays	Television-Weekends	Radio Weekdays	Radio Weekends	Newspapers	Magazines
Television-weekdays	1.00	.59*	.23*	.08	.06	.04
Television-weekends		1.00	.004	.04	.02	.003
Radio-weekdays			1.00	.75*	.02	.09
Radio-weekends				1.00	.05	.12
Newspaper					1.00	.51*
Magazines						1.00

*p < .05

Table VIII

Results of Comparing Mass Media Exposure to
Income and Social Class

Type of Media	Income		Social Class	
	x ²	Level of Significance	x ²	Level of Significance
Television Weekday	23.53	.001	5.03	.100
Television Weekend	38.14	.001	2.26	.500
Radio Weekday	15.48	.010	4.73	.100
Radio Weekend	12.36	.010	4.33	.200
Newspapers	9.92	.020	11.95	.010
Magazines	7.89	.050	18.05	.001

parents. As a result senior high respondents may have to give up hours of television exposure to acquire more income.

Print media exposure was found to be significantly related to both income and social class (See Table VIII). As suggested by previous research, adolescents with high incomes and in the upper class spend more time reading newspapers and magazines than other adolescents (Greenburg & Dervin, 1970). It would appear that print media readership depends on income available to buy such media and social status motivation to read such media.

Personal Sources

Family information sources as well as friends and teachers were personal information sources investigated in this study.

Initial Product Information

Mother was the personal information source used most frequently for initial product information (See Table I and Table II). In fact, mother was used more frequently for initial product information than any other information source. It is important to note that friends were the second most frequently used personal source of initial product information. Teachers and strangers were the personal sources used least frequently for initial product information.

Income. Examination of Table III discloses that neither income nor social class are significantly related to utilization of most personal sources of in-

formation for initial product information. The only significant relationship observed for income involves the use of father. The observed relationship indicates increased utilization of father for initial product information by respondents with higher incomes. The examination of junior and senior high school strata show that the relationship is significant for junior high but not senior high adolescents. Possibly junior high respondents increase their use of father as income increases because their father is the source of income. Another significant relationship is observed for income when junior high respondents are investigated. Junior high respondents with zero or high income (\$10 or more weekly) used relatives outside the household more frequently for initial product information than did those junior high respondents who had incomes of less than \$10 a week.

Social Class

Social class was found significantly related to use of friends for initial product information (See Table III). In this case increased levels of social class were found matched with increased frequency of use of friends for initial product information. No other significant relationships were found for social class and initial product information even when junior high and senior high strata were considered.

Subsequent Product Information

As observed for initial product information mother was the personal information source used most frequently for subsequent product information (See Tables IV and V). Similarly, teachers and strangers were the personal sources used least frequently.

Income

Income of respondents was found significantly related to frequency of use of mother and friends for subsequent product information. As found by Ward and Robertson (1970) adolescents with low or medium income utilized mother more frequently as a source of subsequent product information than respondents who earned no income or high income.

The relationship between income and the use of father for subsequent product information reaches the .05 level of significance in the case of junior high respondents. The observed relationship is in agreement with that previously discussed for use of father for initial product information. Junior high respondents earning a high amount of money utilized father more frequently for subsequent product information than those who earned less money. The other significant relationship observed between income and subsequent use of personal sources was found for friends. Friends were used more for subsequent product information by respondents with high incomes than by those who earned less money.

Separate consideration of junior high and senior high respondents resulted in the findings of no significant difference between social class and use of personal information sources for subsequent product information.

The fact that no significant relationships were found between social class and use of parents for product information agrees with the findings of Ward and Robertson (1970).

Decisive Product Information

Mother was considered to be the most important personal information source by more respondents than any other source. Friends was the second most important source of personal information. The primary importance of mother and friends agrees with the high frequency of usage of those sources for initial and subsequent product information. Agreement was also found for the sources considered the least important as strangers and teachers were designated as least important by more respondents than any other personal source.

Income. Income was found to be significantly related to importance attached to different personal information sources ($\chi^2 = 37.41$, level of significance = .001). Again the observed relationships parallel those found for initial and subsequent product information. The higher the income of respondents the more likely it was that respondents considered father and friends to be the most influential sources of personal information. Conversely family members other than father were more important personal information sources to respondents earning low incomes. No other significant relationships were found for income and the importance attached to personal information sources even when junior high and senior high respondents were considered separately.

Social class. The social class of respondents did not affect the importance attached to different personal information sources. As a result adolescents of any social class are more likely to be influenced by mother than any other personal information source. This is true for junior high as well as senior high respondents.

Opinion Leadership

Opinion leadership is an aspect of external search behavior which is applicable to personal information sources. Several aspects related to opinion leadership are examined in this study: (1) the relationship of opinion leadership to income and to social class, (2) the relationship of opinion leadership to mass media exposure, and (3) the relationship of opinion leadership to the relative age and family income of the recipient of the interpersonal communication.

Analysis of the relationship of the opinion leadership score, derived from a set of self-designation items, to income using Chi-square analysis indicates that opinion leaders are more likely to have high amount of earnings than lower earnings (See Table IX). As a result, it would appear that the amount of income earned by adolescents affects their ability to influence the consumption behavior of others. This finding is in agreement with those of Summers (1970) and Corey (1971).

Table IX

Results Comparing Opinion Leadership
with Income and Social Class

<u>Variables</u>	<u>χ^2</u>	<u>Level of Significance</u>
Income	11.29	.02
Social Class	3.51	.50

Relating opinion leadership to social class reveals that adolescent opinion

leaders seem to be found in approximately equal proportions at all strata of society (See Table IX). As in the case of income and opinion leadership, the relationship found between social class and opinion leadership is consistent with previous research of Katz and Lazarsfeld (1955). Alternately, these findings are contrary to those reported by Summers (1970), Corey (1971) and Fenton and Leggett (1971). They found that opinion leaders were more likely to be in the higher social classes than nonopinion leaders.

Mass media exposure. In order to test the relationship between opinion leadership and amount of mass media exposure, Chi-square analysis was employed.

Except for weekend television viewing, no significant relationships were found between opinion leadership and amount of exposure to television and radio (See Table X). The significant finding for amount of exposure to television on weekends is contrary to that reported by Katz and Lazarsfeld. Self-designated opinion leaders spent fewer hours watching television on weekends than nonopinion leaders. Possibly, adolescents who consider themselves opinion leaders spend their spare time on weekends influencing friends rather than watching television. Unfortunately, this reasoning does not explain why non-opinion leaders would watch more television on weekends if they are the ones who are being influenced.

Table X

Opinion Leadership and Exposure to Mass Media

Type of Media	χ^2	Level of Significance
Television Weekday	1.13	.30
Television Weekend	5.93	.02
Radio Weekday	3.54	.10
Radio Weekend	3.78	.10
Newspapers	5.65	.02
Magazines	13.14	.001

Conversely, the significant relationships between opinion leadership and amount of exposure to newspapers and magazines do agree with those reported by Katz and Lazarsfeld. Respondents who considered themselves opinion leaders spent more time reading newspapers and magazines than those who did not consider themselves opinion leaders. Newspapers and magazines appear to be a more important source of information for adolescent opinion leaders than television and radio.

Relative age and family income. Reynolds and Myers (1966) suggested that the older the adolescent the higher the status accorded him by younger adolescents and the more likely it is that he is able to influence their attitudes. Similarly, previous studies by Summers (1970) and Corey (1971) have reported positive relationships between the incidence of opinion leadership and income.

Results from the present study provide support for the idea that opinion leadership is a horizontal phenomenon within groups of adolescents who are similar ages and are in families with the same incomes. As shown in Table VI, the majority of respondents felt that influential friends were the same age (88 percent) and that the incomes (78 percent) of their families were comparable.

Marketer-Dominated Sources

Initial Product Information

Store visits were used more frequently for initial product information than other marketer-dominated sources (See Table I and Table II). The least frequently used marketer-dominated source is shopkeepers or clerk.

Table XI

Relative Age and Family Income
of Influential Friends

Relative Age of Friends*	Observed Frequencies	Expected Frequencies
Younger	24	8
The Same	717	8
Older	75	800
Total	<u>816</u>	<u>816</u>
<hr/>		
Relative Family Income of Friends*	Observed Frequencies	Expected Frequencies
Less than	85	8
The Same	635	8
More than	96	800
Total	<u>816</u>	<u>816</u>

*The observed distribution was found to be significantly different from the expected distribution at the .001 level.

Income. As shown in Table III income did not affect frequency of use of marketer-dominated sources for initial product information. Neither junior nor senior high adolescents with different incomes varied significantly in their frequency of use of marketer-dominated sources for initial product information.

Social class. Further examination of Table III results in the conclusion that social class affects use of store visits for initial product information.

A positive relationship was observed between social class and use of store visits as upper class respondents used store visits more frequently than those in other social classes. No other effects of social class were observed in the case of use of marketer-dominated sources.

Subsequent Product Information

Store visits were utilized more frequently for subsequent product information than any other marketer-dominated source. As shown in Table IV shopkeepers or clerks were the least frequently utilized marketer-dominated source.

Income. Income was significantly related to the use of two marketer-dominated information sources, shopkeepers or clerks and store visits, for subsequent product information (See Table VI). In both cases the relationship was significant for junior high but not senior high respondents. Income may determine if junior high respondents are able to visit the store and use the clerks in these stores as increased usage of these sources for subsequent product information was coupled with high levels of income.

Social class. The only significant relationship between social class and use of marketer-dominated sources for subsequent product information is for the source of store visits. As encountered for use of store visits for initial product information adolescents in the upper class utilized store visits more frequently for subsequent product information than did those in the middle or lower classes.

Decisive Product Information

Store visits were not only the most frequently used marketer-dominated information sources but also were the most influential. As experienced for mass media sources the least used marketer-dominated source, shopkeepers or clerks, were not the least influential. Mail catalogs rather than shopkeepers or clerks were assigned the role of least important marketer-dominated source.

Income

No significant relationships were found in the importance assigned to the different marketer-dominated information sources and income of respondents. As a result the increased use of store visits for subsequent product information by respondents with high income does not mean that these earners of high incomes are more influenced by store visits than those earning lower incomes.

Social class. Significant differences exist between social class and marketer-dominated information sources considered the most important by senior high respondents. Senior high respondents in the upper class were more likely to consider store visits as the most important marketer-dominated source of product information.

Brand Identification and Slogan Recall

Another aspect of external search processes involving marketer-dominated sources investigated in this study is brand identification and slogan recall. In order to measure these two aspects of brand awareness, respondents were asked to associate 20 brand names with generic product categories and to identify the sponsor of 10 different advertising slogans. The brand names and advertising slogans used were selected for their variety and recency of ap-

pearance in the media or retail stores.

The total number of brands correctly identified and slogans correctly recalled were determined for each respondent. The results of statistically relating these two variables to income and social class are shown in Table XII.

Table XII
Results Comparing Brand Awareness
with Income and Social Class

Type of Brand Awareness	INCOME		SOCIAL CLASS	
	χ^2	Level of Significance	χ^2	Level of Significance
Brand Identification	34.25	.001	32.80	.001
Slogan Recall	6.31	.800	6.58	.500

Brand identification is statistically related to income and social class. Although income and social class variables have been found to be statistically related to slogan recall of adults, present research results do not support a similar conclusion for adolescents.

For both income and social class, rising levels of the variables were found to be positively related to brand identification. Adolescents with high earnings and in the upper class were able to identify more brands than other adolescents. Evidently, increased income and social status are linked to increased external search behavior. It may be that the other roles of promotion--liking, conviction, etc.--should be emphasized when aiming at adolescents with high income and social status. The significant relationship observed between social status and brand awareness supports that found for adolescents by Guest (1942) three decades ago.

Retail Patronage Motives

All marketers regardless of their position in the channel are concerned with external search behavior of consumers. This is especially true of the retailer since he is the first to feel the effect of a consumer's decision. The retailer has many ways to send messages to the potential consumer. A substantial portion of his messages are sent from his retail store and will not be received unless the consumer visits his store. As a result of the importance of store visitation patterns of consumers to the retailer, the reasons for adolescent store choice were investigated.

The results of statistically examining the relationship between income and social class and retail store patronage motives with the use of Friedman two-way analysis of variance are presented in Table XIII. Neither income nor social class affect the importance ascribed to these different motives.

Subsequent Friedman analysis on the differences in the ranks of the different motives showed that significant differences exist between the impor-

portance respondents attach to different motives. Quality is considered the most important reason for patronizing a store by more respondents than any other motive. As shown in Table XIV assortment and style were also considered to be relatively important factors in the store selection decision. Conversely, respondents considered friendliness and convenience as relatively unimportant reasons for choosing a store. Interestingly, the low importance attached to friendliness and convenience aspects is contrary to behavior normally ascribed to low income earners and lower class buyers. The behavior observed for adolescents is more like that of adults with high amounts of income and in the upper social class. The relatively discretionary nature of the income of adolescents may explain why they are more concerned with the quality of store patronized rather than their price or friendliness of personnel.

Table XIII

Results Comparing Retail Store Patronage
Motives with Income and Social Class

	Friedman Statistic	Level of Significance
Income	.26	.98
Social Class	.75	.80

Table XIV

Relative Importance of Retail Store
Patronage Motives

Motive	Average (mean) Rank
Quality	2.82
Assortment	3.27
Style	4.04
Trustworthiness	4.58
Service	4.78
Convenience	5.10
Prices	5.20
Friendliness	5.65

Summary

The purpose of this study was to examine the effect of social class and income on the external search behavior of adolescents. The aspects of external search behavior examined were: (1) use of mass media information sources; (2) use of interpersonal information sources; and, (3) use of marketer-dominated sources.

Data utilized in this analysis were collected through the use of a questionnaire which was administered during March, 1971 to 1,300 adolescents attending school in the Columbus, Ohio public school system. Of this total,

1,280 usable questionnaires were obtained. The analyses were performed with the use of the Chi-square test, Wilcoxon matched-pairs signed-ranks test, and Friedman two-way analysis of variance. For all statistical tests, an alpha level of .05 or lower was used.

Summary of Findings

For mass media information sources---

1. Television was the mass media source used most frequently for "initial", "subsequent", and "decisive" product information.
2. Income and social class were not significantly related to adolescents utilization of mass media sources of product information.
3. Exposure to mass media was related to income but not to the social class of the respondents.

For interpersonal sources of information---

1. The respondent's mother was the most frequently utilized interpersonal source of information.
2. Income but not social class was significantly related to adolescents utilization of inter-personal sources of product information for "initial", "subsequent", and "decisive" product information.
3. Income and exposure to print media were related to opinion leadership.
4. Opinion leaders were found to be similar to non-opinion leaders with regard to their age and family income.

For marketer-dominated sources of information---

1. Store visits were the most frequently used and most influential marketer-dominated sources of product information.
2. Income was significantly related to use of marketer-dominated sources for "subsequent", but not "initial" or "decisive" product information while social class was significantly related to all three dimensions of marketer-dominated sources of product information.
3. Brand identification, but not slogan awareness, was related to both income and social class of adolescents.
4. Adolescents' patronage motives did not vary as a result of income or social class.

Areas for Future Research

External search behavior of adolescents should be subjected to further investigation. A need for further research stems from the requirements of marketers for a better understanding of the consumer behavior of adolescents.

This need for further research encompasses two broad areas in the consumption decision-making process of adolescents. Future research should further investigate the nature of adolescents' external search behavior as well as the behavior that follows external search behavior.

Increased information needs about the nature of external search behavior include the reasons why this behavior occurs, how the search behavior patterns

of adolescents change as they mature, how external search behavior of adolescents in different geographical locations compare, and the type of consumption-related information that adolescents acquire from each information source.

In addition to further research on the nature of external search behavior of adolescents, future research should focus upon adolescent consumption behavior after external search has taken place. The most pressing need is to link aspects of external search behavior to subsequent consumption decisions. The usefulness of the information provided by this research effort will be enhanced if the different aspects of external search behavior are related to aspects of buyer behavior, such as brands purchased, product usage rates, repurchase patterns, stores patronized, and prices paid. The issue is whether adolescents with different external search behavior have different consumption behavior.

Footnote

1. Both authors are assistant Professors of Marketing at their respective schools.

References

- Bowerman, Charles E. & Kinch, John W. Changes in Family and Peer Orientation of Children Between Fourth and Tenth Grades. Social Forces, 1959, 37, 206-211.
- Coleman, James S. The Adolescent Society. New York: Free Press of Glencoe, 1961.
- Collazzo, Charles J., Jr. Effects of Income Upon Shopping Attitudes and Frustrations. Journal of Retailing, 1966, 42, 1-7.
- Corey, Lawrence G. People Who Claim to be Opinion Leaders: Identifying Their Characteristics by Self Report. Journal of Marketing, 1971, 35, 48-53.
- Douvan, Elizabeth & Adelson, John. The Adolescent Experience. New York: John Wiley and Sons, Inc., 1966.
- Engel, James F.; Kollat, David T.; & Blackwell, Roger D. Consumer Behavior. New York: Holt, Rinehart, and Winston, Inc., 1968.
- Fenton, James S. & Leggett, Thomas R. A Way to Find Opinion Leaders. Journal of Advertising Research, 1971, 11, 21-25.
- Gilkinson, Paul. What Influences the Buying Decisions of Teenagers. Journal of Retailing, 1965, 41, 31-41, 48.
- Greenburg, Bradley, & Dervin, Brenda. Mass Communications Among the Poor. Public Opinion Quarterly, 1970, 34, 224-235.
- Guest, Lester P. Brand Loyalty: Twelve Years Later. Journal of Applied Psychology, 1955, 39, 405-408.
- _____. The Genesis of Brand Awareness. Journal of Applied Psychology, 1942, 26, 800-808.
- Howard, John A. & Sheth, J. N. Theory of Buyer Behavior. Changing Marketing Systems. Edited by Reed Moyer. Chicago: American Marketing Association, 1967.
- Katz, Elihu, & Lazarsfeld, Paul F. Personal Influence. New York: Free Press of Glencoe, 1955.
- Katz, Michael, & Rose, Jan. Is Your Slogan Identifiable. Journal of Advertising Research, 1969, 9, 21-26.
- King, Charles W., & Summers, John O. Attitudes and Media Exposure, Journal of Advertising Research, 1971, 11, 26-32.
- Larsen, Carl M. & Wales, Hugh G. Slogan Awareness in the Chicago Market. Journal of Advertising Research, 1970, 10, 38-41.

- McCarthy, E. Jerome. Basic Marketing. 4th ed. Homewood, Ill.: Richard D. Irwin, Inc., 1971.
- Martineau, Pierre. The Pattern of Social Classes. Marketing's Role in Scientific Marketing. Edited by Richard Clewett. Chicago: American Marketing Association, 1957.
- Myers, James H.; Stanton, Roger R.; & Haug, Arne F. Correlates of Buying Behavior: Social Class vs. Income. Journal of Marketing, 1971, 35, 8-15.
- National Industrial Conference Board. A Graphic Guide to Consumer Markets. New York: National Industrial Conference Board, 1969.
- Rainwater, Lee; Coleman, Richard P.; Handel, Gerald. Workingman's Wife. New York: Oceana Publications, Inc. 1959.
- Reynolds, William H., Myers, James H. Marketing and the American Family. Business Topics, 1966, 14, 57-66.
- Rich, Stuart U., & Jain, Subhash C. Social Class and Life Cycle as Predictors of Shopping Behavior, Journal of Marketing Research, 1968, 5, 41-49.
- Robertson, Thomas S. Consumer Behavior, Glenview, Ill.: Scott, Foresman and Company, 1970.
- Siegel, Sidney. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill Book Company, 1956.
- Slocum, John W., Jr., & Mathews, H. Lee. Social Class and Income as Indicators of Consumer Credit Behavior. Journal of Marketing, 1970, 34, 71-78.
- Solomon, Daniel. Influences on the Decisions of Adolescents. Human Relations, 1963, 16, 45-60.
- Summers, John O. The Identity of Women's Clothing Fashion Opinion Leaders. Journal of Marketing Research. 1970, 7, 178-185.
- Teen-Age Consumer: Dynamic Force in the U. S. Economy, Senior Scholastic, February 25, 1965, p. 5.
- The Los Angeles Chamber of Commerce. The Dynamics of the Youth Explosion: A Look Ahead. Los Angeles: Los Angeles Chamber of Commerce, 1967.
- Thompson, Bryan. An Analysis of Supermarket Shopping Habits in Worcester, Massachusetts. Journal of Retailing, 1967, 43, 17-29.
- Walters, C. Glenn & Paul, Gordon W. Consumer Behavior: An Integrated Framework. Homewood, Ill.: Richard D. Irwin, Inc., 1970.
- Ward, Scott, & Robertson, Thomas S. Family Influences on Adolescent Consumer Behavior. Paper presented to the Association for Consumer Research, Amherst, Mass., August, 1970.
- Wasson, Chester R. Is it Time to Quit Thinking of Income Classes? Journal of Marketing, 1969, 33, 54-69.
- White, R. Clyde. Social Class Differences in the Uses of Leisure. American Journal of Sociology, 1955, 61, 145-150.
- "Youth Appeal...A Creative Booby Trap?" Grey Matter, 1971, 17.

OPINION LEADERSHIP AND INNOVATIVENESS: OVERLAP AND VALIDITY

Jacob Jacoby
Purdue University

The construct of opinion leadership has played an important cross-disciplinary role in studies dealing with the adoption and diffusion of products, services, and ideas. Originally described by Lazarsfeld, Berelson, and Gaudet (1948) in their study of the 1940 presidential election, the construct has since been employed by rural sociologists, agricultural economists, communication theorists, social psychologists, and marketers, among others, to study a variety of interpersonal communication phenomena.

While there are various specific measures of opinion leadership, they can all generally trace their origin to one of three different techniques. The self-designating technique is one where the individual is asked to indicate how much of an opinion leader he perceives himself to be. The sociometric approach is one in which all members of a given group are asked to identify those group members considered most influential with respect to the object or idea under consideration. The key informant technique involves first identifying a limited number of people within the group assumed to be knowledgeable regarding the patterns of influence within that group, and then asking them to identify the influentials in that group.

Given the importance and high frequency with which the term opinion leadership appears in the consumer behavior literature, it is surprising to see so few data presented which bear on the interrelationship of these various measurement techniques and on the construct validity of opinion leadership itself. Indeed, the literature seems to contain only two studies which cite original data bearing on the interrelationship of two or all three of these techniques. Katz and Lazarsfeld (1955, pp. 137-161) provide percentage-based evidence to indicate that a high degree of agreement exists between self-designating and sociometric approaches. Rogers and Catarno (1962, p. 441) cite data from an unpublished doctoral dissertation in which all three methods of measuring opinion leadership were applied to 28 dairy farmers. "Self-designating opinion leadership scores were correlated .300 with the number of sociometric choices, and .640 with composite opinion leadership rating by four key informants. The number of sociometric choices was correlated .876 with the composite key informants ratings." However, while data based on assessing opinion leadership with respect to one object or idea do provide an indication of convergent validity, they provide no indication of discriminant validity, and it is the latter which is essential for establishing construct validity (cf. Campbell and Fiske, 1959).

Accordingly, it was the primary purpose of this investigation to apply all three techniques for assessing opinion leadership to several groups of subjects, and to measure such opinion leadership across three different areas of possible influence, so as to assess the convergent and discriminant validity of this construct. This is essentially the multimethod-multitrait approach proposed by Campbell and Fiske (1959). In this regard, construct validity "is primarily concerned 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. We believe that before one can test the relationships between a specific trait and other traits, one must have confidence in one's measure of that trait (1959, p. 100)."

Other purposes of this investigation were: (a) to examine the question of opinion leadership overlap across product categories (cf. King and Summers, 1970); (b) to provide another test of the relationship between opinion leadership and innovativeness (cf. Robertson, 1970, p. 137); and (c) to examine the between-method consistency in identifying innovativeness using the key informant and sociometric techniques.

Method

Subjects

Application of the key informant and sociometric techniques is meaningless if all group members do not know, or know of, one another. Consequently, groups which were relatively small and cohesive had to be utilized. College fraternities, sororities, and their pledge-classes satisfy these criteria. Thus, the subjects in this investigation were two fraternity pledge-classes ($n_1 = 13$; $n_2 = 27$), and three entire sororities ($n_3 = 60$; $n_4 = 70$; $n_5 = 70$)¹ at Purdue University during the Spring 1971 semester. The response rate was 100% for Groups 1 and 2; 67% (40/60) for Group 3; 54% (38/70) for Group 4; and substantially less than 50% for Group 5. Accordingly, the data for Group 5 were discarded.

Traits

To assess discriminant validity using the multimethod-multitrait approach, "more than one trait as well as more than one method must be employed in the validity process" (Campbell and Fiske, 1959, p. 81). Accordingly, the two groups of males were examined for opinion leadership in three different product categories: clothing, alcoholic beverages, and LP records. The trait categories for the two groups of females were: clothing, cosmetics, and room decorations.

Instruments

Two slightly different self-designating scales were used for the fraternities and sororities. The scale for the fraternity pledge classes was a modified version of the Rogers and Catarno (1962, pp. 439-440) self-designating opinion leadership scale. The modification attempted to increase reliability and sensitivity by providing more response alternatives for the six questions than were provided in the original scale. A copy of this scale for clothing is included in Appendix A. In all, each fraternity pledge responded to 18 questions--six for each product (i.e., trait). The self-designating scale for the sorority members consisted of seven-item scales for each product patterned after King and Summers (1970, p. 45), and also modified to increase reliability and sensitivity. An example is provided in Appendix A.

The sociometric technique involved asking each member of the group two questions for each product category, and then determining how each member of the group was ranked on these items by his peers. For example,

the fraternity pledges were given the following two questions for clothing: "If you were going to a semi-formal dance, which fellow pledges would you go to to get ideas on what to wear?" and "Your parents just gave you some money to buy some new summer clothes. Which fellow pledges would you want to help you select them?" In responding to these questions, each subject rank ordered the five "best" sources he knew of within his group. An index was derived for each subject within each group based upon the average rank he was assigned by his fellow group members. This index was continuous in that, while subjects who never received any votes were always scored as 1.00 and subjects who were always ranked as the "best" source received a score of 6.00, all other combination of rankings resulted in some score (usually not a whole integer) between 1.00 and 6.00.

The key informants in the two sororities were the two presidents and two social chairmen. The key informants in the fraternity pledge-classes were selected on the basis of their sociometric scores. The format for the key informant instrument was comparable to that of the sociometric instrument. The basic difference between the two was that the questions on the sociometric instrument were phrased in the first person ("Whom would you go to when . . . ?"), while those on the key informant instruments were phrased in the third person ("Whom would the majority of the other members go to when . . . ?"). Again, an index was derived for each person for each product category based on his rank on the two questions relating to that product category.

Finally, in the spirit of exploratory investigation, the sociometric and key informant instruments for Groups 3 and 4 also included items designed to assess innovativeness. While the self-designating approach is coming into frequent use as a measure of innovativeness, the sociometric and key informant techniques do not appear among the approaches used to assess innovativeness (cf. King and Ryan, 1971). Therefore, a single question for each of the three product categories was included as a crude measure of innovativeness on both the sociometric and key informant instruments. For example, all participants respond to the question "If you want information about a new cosmetic product, which three girls in the house would you ask?" by citing their best, second best, and third best sources. An innovativeness index was derived for each subject based upon the average rank she was assigned by her sorority sisters. Similarly, the key informants received a single question for each product category--e.g., "Overall, which three girls in the house would be most generally considered to always have the latest look in make-up?"--and responded by providing whom she thought would be the first, second, and third best sources. Again, an innovativeness index was derived.

Results

The multimethod-multitrait matrices for Groups 1-4 are presented in Tables 1-4, respectively. Campbell and Fiske (1959) suggested that one should examine convergent validity before proceeding to assess discriminant validity. The former is accomplished through examining the correlation coefficients along the validity diagonals. These values "should be significantly different from zero and sufficiently large to encourage further examination of validity (p. 82)."

Table 1. The multi-method x multi-trait matrix for Group 1 (n = 13). (N.B. The validity diagonals are the three sets of unenclosed coefficients. Each heterotrait-monomethod triangle is enclosed by a solid line. Each heterotrait-heteromethod triangle is enclosed by a broken line.) Minimal coefficient values necessary to reach statistical significance (11 df, one-tailed test) are:

P < .05 r = .476
 p < .01 r = .634

	Self Designating			Sociometric			Key Informant		
	Alcohol	Records	Clothing	Alcohol	Records	Clothing	Alcohol	Records	Clothing
<u>Self Designating</u>									
Records		.065							
Clothing		-.021	.118						
<u>Sociometric</u>									
Alcohol		.759	.127	.247					
Records		.129	.394	.449	.283				
Clothing		.163	.443	.536	.153	.804			
<u>Key Informant</u>									
Alcohol		.779	.114	.248	.965	.261	.083		
Records		-.087	.166	.177	.196	.696	.448	.211	
Clothing		.325	-.038	.545	.196	.519	.622	.306	.354

Table 2. The multi-method x multi-trait matrix for Group 2 (n = 27). (N.B. The validity diagonals are the three sets of unenclosed coefficients. Each heterotrait-monomethod triangle is enclosed by a solid line.) Minimal coefficient values necessary to reach statistical significance (df = 25, one-tailed test) are:

p < .05 r = .323
 p < .01 r = .445

	Self Designating			Sociometric			Key Informant		
	Alcohol	Records	Clothing	Alcohol	Records	Clothing	Alcohol	Records	Clothing
<u>Self Designating</u>									
Records		.225							
Clothing		.589	.138						
<u>Sociometric</u>									
Alcohol		.591	.501	.320	.248	.524			
Records	.754	.319	.594						
Clothing	.194	.556	.165						
<u>Key Informant</u>									
Alcohol		.493	.235	.515	.474	.462	.457	.721	.401
Records	.298	.051	.251	.289	.187	.281			
Clothing	.511	.372	.358	.552	.352	.432			

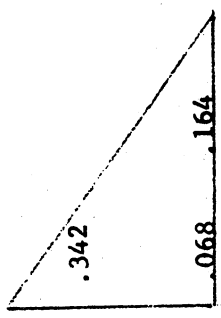
Table 3. The multi-method x multi-trait matrix for Group 3 (n = 40). (N.B. The validity diagonals are the three sets of unenclosed coefficients. Each heterotrait-monomethod triangle is enclosed by a solid line. Each heterotrait-heteromethod triangle is enclosed by a broken line.)

Minimal coefficient values necessary to reach statistical significance (df = 38, one-tailed test) are:

p < .05 r = .264
 p < .01 r = .363

	Self Designating			Sociometric			Key Informant		
	Clothes	Cosmetics	Room Dec's	Clothes	Cosmetics	Room Dec's	Clothes	Cosmetics	Room Dec's

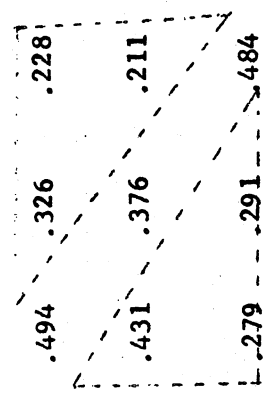
Self Designating



Cosmetics

Room Dec's

Sociometric

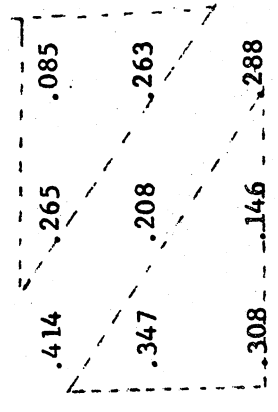


Clothes

Cosmetics

Room Dec's

Key Informant



Clothes

Cosmetics

Room Dec's

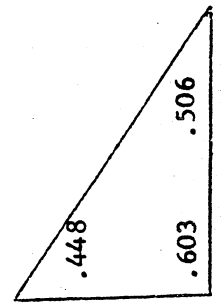
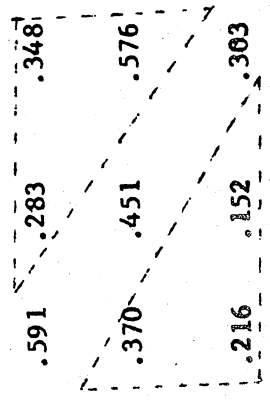
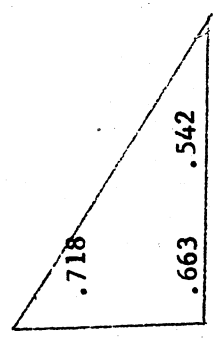


Table 4. The multi-method x multi-trait matrix for Group 4 (n = 38). (N.B. The validity diagonals are the three sets of unenclosed coefficients. Each heterotrait-monomethod triangle is enclosed by a solid line. Each heterotrait-heteromethod triangle is enclosed by a broken line.)

Minimal coefficient values necessary to reach statistical significance (df = 36, one-tailed test) are:

$p < .05$ $r = .272$
 $p < .01$ $r = .377$

	Self Designating			Sociometric			Key Informant		
	Clothes	Cosmetics	Room Dec's	Clothes	Cosmetics	Room Dec's	Clothes	Cosmetics	Room Dec's
<u>Self Designating</u>									
Cosmetics		.421							
Room Dec's		.411	.261						
<u>Sociometric</u>									
Clothes	.421	.043	.023	.776					
Cosmetics	.289	.043	-.041	.098	.160				
Room Dec's	.265	.112	.449						
<u>Key Informant</u>									
Clothes	.373	.030	.017	.754	.994	.167			
Cosmetics	.308	.053	.015	.675	.889	.378	.882		
Room Dec's	.221	-.022	.419	.001	.068	.589	-.011	.055	

The nine validity diagonal coefficients of Table 1 range from .166 to .965, with a median of .622. Seven of these values are statistically significant at $p < .05$ or better. The validity diagonal coefficients of Table 2 range from .051 to .591, with a median value of .358. Six of these are significant at $p < .05$ or better, and a seventh nearly so ($p \approx .06$). Table 3 contains validity diagonal coefficients which range from .208 to .591, with a median value of .414. Eight of these coefficients are statistically significant at $p < .05$ or better. Finally, the validity diagonal coefficients of Table 4 range from .043 to .889, with a median value of .421. Seven of these coefficients are statistically significant at $p < .02$ or better.

Thus, approximately 80 % of these coefficients (28 out of 36) reach acceptable levels of statistical significance and a 29th nearly so ($p \approx .06$). The median value across all 36 coefficients is .440. These values are surprisingly high in view of the limits that the reliability of an instrument places on its validity. Rogers and Catarno (1962, p. 446) reported a split-half reliability of .703 for their self-designating scale. While the reliabilities of the specific instruments used in the investigation were not established, it is very probable that they were not much higher. It would appear, therefore, that convergent validity has been adequately established.

Three criteria must be satisfied in order to demonstrate discriminant validity. First, "a validity value for a variable should be higher than the correlation obtained between that variable and any other variable having neither trait nor method in common (Campbell & Fiske, 1959, p. 82)." This requires that each validity diagonal value be higher than the value lying in its column and row in the heteromethod-heterotrait triangles. For example, the validity diagonal coefficient based on the Self Designating and Key Informant assessment of clothing opinion leadership in Table 1 is .545. It is higher than all eight of the heteromethod-heterotrait coefficients lying in its column (.247, .449, .248, .177) and row (.335, .038, -.196, .519). Each of the nine validity diagonal coefficients was involved in eight direct comparisons. Across all groups, approximately 72% (207/288) of these comparisons satisfied the first discriminant validity criterion. More specifically, 65/72 did so in Group 1, 26/72 in Group 2, 56/72 in Group 3, and 60/72 in Group 4.

The second criterion for establishing discriminant validity requires "that a variable correlate higher with an independent effort to measure the same trait than with measures designed to get at different traits which happen to employ the same method (Campbell and Fiske, 1959, p. 83)." This requires that each value along the validity diagonals be compared against the two values in each of the two heterotrait-monomethod triangles which use the same two methods, but which depict how well that trait correlates with other traits using the same methods. As an example, consider the first validity diagonal value in Table 1, $r = .759$. This coefficient represents the correlation between two methods (self-designating and sociometric) of assessing the same trait (opinion leadership with respect to alcoholic beverages). It must, therefore, be demonstrated that this value is higher than correlations between that trait and the other two traits using either of the two methods in question. This criterion is completely satisfied in this instance inasmuch as the correlations between being an alcoholic beverage opinion leader and either a record or clothing opinion leader are only .065 and =.021, respectively, for the self-designating monomethod-heterotrait block, and only .283 and .153, respectively, in the sociometric monomethod-heterotrait block.

There are 36 comparisons (9 validity diagonal coefficients X 2 values in each of the appropriate monomethod-heterotrait blocks) to be made for each of Tables 1-4. Thirty such comparisons in Table 1 satisfied this second criterion, 15 in Table 2, 13 in Table 3, and 22 in Table 4. Overall, 80 of 144 comparisons (55.44%) satisfied this second criterion. These results are considerably better than most such results reported in the behavioral science literature.

The third and final criterion necessary for providing evidence of discriminant validity "is that the same pattern of trait interrelationships be shown in all of the heterotrait triangles of both the monomethod and heteromethod blocks (Campbell and Fiske, 1959, p. 83)". Accordingly, the three coefficients in each of the nine triangles were rank ordered for each Group, and Kendall's coefficient of concordance (W) was computed to determine the extent of agreement among the nine sets of ranks within each Group (Siegel, 1956, pp. 229-238). Tables 5-8 present these data. In summary, the values of S for Group 1 (78) and Group 4 (131) were significant at $p < .01$, but failed to reach traditional levels of statistical significance in Groups 2 and 3.

Table 5

Kendall's Coefficient of Concordance (W) for Group 1

Triangle	Alcoholic Beverages and Records	Alcoholic Beverages and Clothing	Records and Clothing
1	2	3	1
2	2	3	1
3	2	3	1
4	3	2	1
5	2	3	1
6	2	3	1
7	3	2	1
8	2	3	1
9	2	1	2
	$\Sigma = 21$	23	10
$S = \frac{(21 - 18)^2 + (23 - 18)^2 + (10 - 18)^2}{12(78)} = 78$			
$W = \frac{78}{18(27 - 3)} = .481$			

Table 6

Kendall's Coefficient of Corcordance (W) for Group 2

Triangle	Alcoholic Beverages and Records	Alcoholic Beverages and Clothing	Records and Clothing
1	2	1	3
2	2	3	1
3	2	1	3
4	1	3	2
5	3	1	2
6	3	1	2
7	3	2	1
8	1	2	3
9	<u>3</u>	<u>1</u>	<u>2</u>
	$\Sigma = 20$	15	19
$S = \frac{(20 - 18)^2 + (15 - 18)^2 + (19 - 18)^2}{12(14)} = 14$			
$W = \frac{14}{18(27 - 3)} = .084$			

Table 7

Kendall's Coefficient of Concordance (W) for Group 3

Triangle	Clothing and Cosmetics	Clothing and Room Decorations	Cosmetics and Room Decorations
1	1	3	2
2	1	2	3
3	3	1	2
4	1	3	2
5	1	2	3
6	1	2	3
7	1	2	3
8	1	2	3
9	<u>3</u>	<u>2</u>	<u>1</u>
	$\Sigma = 13$	19	22
$S = \frac{(13 - 18)^2 + (19 - 18)^2 + (22 - 18)^2}{12(42)} = 42$			
$W = \frac{42}{18(27 - 3)} = .259$			

Table 8

Kendall's Coefficient of Concordance (W) for Group 4

Triangle	Clothing and Cosmetics	Clothing and Room Decorations	Cosmetics and Room Decorations
1	1	2	3
2	1	3	2
3	1	3	2
4	1	2	3
5	1	3	2
6	1	2	3
7	1	2	3
8	1	2	3
9	<u>1</u>	<u>3</u>	<u>2</u>
	$\Sigma = 9$	22	23
$S = (9 - 18)^2 + (22 - 18)^2 + (23 - 18)^2 = 122$			
$W = \frac{12(122)}{18(27 - 3)} = .753$			

Table 9 summarizes the results of the four validity tests for each of the four groups. It appears as if construct validity was adequately established in Groups 1 and 4 but not in Groups 2 and 3.

Table 9

Summary of Tests for Establishing Construct Validity of Opinion Leadership

	Group 1	Group 2	Group 3	Group 4
<u>Convergent Validity</u>				
Median r	.622	.358	.414	.421
Number Significant	7/9	6/9	8/9	7/9
<u>Discriminant Validity</u>				
Ratio of tests satisfying criterion 1	65/72	26/72	56/72	60/72
Ratio of tests satisfying criterion 2	30/36	15/36	13/36	22/36
Consistency of trait inter- relationship patterns (criterion 3)	W=.481	W=.084	W=.259	W=.753
	p <.01	n.s.	n.s.	p <.01

Examination of the monomethod-heterotrait triangles of Tables 1-4 reveals that, overall, these coefficients tend to be positive and significant. With respect to Group 1, seven of nine values are positive and one of these is significant. All nine coefficients of Table 2 are positive and five of these are significant. Similarly, all nine coefficients of Table 3 are positive and seven of these are statistically significant. Finally, eight of the nine coefficients in Table 4 are positive of which four are significant. Overall, 33 of 36 coefficients were positive and nearly half (17/36) were significant.

There are several implications which stem from these immediately preceding data. One such implication is that they reflect strong methods variance. A second implication is that these data offer moderate to strong support for the contention that opinion leadership overlap exists across product categories (cf. King and Summers, 1970; Marcus and Bauer, 1964). Moreover, the more similar the product categories, the greater the degree of overlap. As examples, there is greater overlap within the sororities between clothing and cosmetic opinion leaders than between clothing and room decoration opinion leaders, or between cosmetic and room decoration opinion leaders (cf. Tables 7 and 8). This is the same pattern of results obtained in Myers and Robertson's (1972) study of 246 Los Angeles housewives. While none of the three products used with the fraternity pledges are similar to each other, it is still obvious that there is less opinion leadership overlap between alcoholic beverage and record opinion leaders than between alcoholic beverage and clothing or clothing and record opinion leaders (cf. Tables 5 and 6).

A third, and probably more significant, implication to be derived from these data is that, to the extent that there is opinion leader overlap, the three "traits" selected to assess the validity of opinion leadership as a construct (i.e., opinion leadership in three specific product categories) are not truly independent and, as such, provide a more stringent test of construct validity. That is, had the sociometric, self-designating, and key informant methods been used to assess opinion leadership (for clothing), brand loyalty (in dentifrices), and deal proneness (in detergent) rather than having assessed opinion leadership for clothing, opinion leadership for cosmetics, and opinion leadership for room decorations, it would probably have been easier to establish validity for the construct of opinion leadership. Under these circumstances, the degree of construct validity actually obtained must be considered very encouraging.

Table 10 contains the intercorrelations among the three methods of assessing overall opinion leadership for all four groups. These coefficients are based upon "general opinion leadership" scores developed from summing each subject's separate opinion leadership scores across all three product categories. The fact that these coefficients are considerably higher than when opinion leadership was considered on a product-by-product basis (median value of .605 versus .440) also tends to support the notion of opinion leadership overlap. The values in Table 10 would have been considerably lower had opinion leadership correlated either negatively or negligibly across product categories.

Table 10

Intercorrelation of Methods Using Overall Opinion Leadership Scores
(Values based on cumulating scores across all three product categories)

	Group 1 (n = 13)		Group 2 (n = 27)	
	Self-designating	Sociometric	Self-designating	Sociometric
Sociometric	.776***		.869***	
Key Informant	.500***	.821***	.572***	.637***
	Group 3 (n = 40)		Group 4 (n = 38)	
	Self-designating	Sociometric	Self-designating	Sociometric
Sociometric	.552***		.317**	
Key Informant	.471***	.778***	.268*	.856***

*p <.06

**p <.05

***p <.01

Table 11 presents the intercorrelations between the opinion leadership index and the single item innovativeness scores for the three products employed with the sorority women. These data add support to the findings of others (cf. Robertson, 1970, pp. 136-137; Myers and Robertson, 1971; Rogers and Stanfield, 1968; Tigert and Arnold, 1971) that a positive relationship exists between opinion leadership and innovativeness. Indeed, the coefficients seem incredibly high (median = .800). However, the fact that they are based upon only a single question dictates that they be interpreted with extreme caution.

Table 11

Correlations Between "Innovativeness" and "Opinion Leadership" Scores

	Key Informant		Sociometric	
	Group 3 (n = 40)	Group 4 (n = 38)	Group 3 (n = 40)	Group 4 (n = 38)
Clothes	.208	.814*	.878*	.597*
Cosmetics	.785*	1.000*	.720*	.979*
Room Decorations	.683*	.998*	.452*	.926*

*p <.01

Table 12 presents the multimethod-multitrait matrices for the single item innovation indices contained within the sociometric and key informant instruments administered to Groups 3 and 4. Both matrices provide adequate evidence of convergent validity. The six coefficients comprising the two validity diagonals range from .509 to .851, and all are significant at $p < .01$. Twenty-three of twenty-four tests satisfy the first criterion of discriminant validity, viz., that the values on the validity diagonal be higher than the coefficients in their respective rows and columns. The only exception occurs in regard to assessing clothing innovativeness in Group 4. The second criterion (as described above) requires that 12 comparisons be made for each of these two matrices. All 12 comparisons of Group 3 satisfy this criterion while 11 of the 12 comparisons do so in Group 4. The third discriminant validity criterion requires that the same pattern of relationships be manifested within all triangles. A precisely identical pattern is manifested in all triangles of Group 4, while Group 3 manifests only one minor inversion: the clothing-room decoration and cosmetics-room decoration correlation coefficients exchange second and third place ranks in one of the four triangles. While the tables for critical values of S in the Kendall coefficient of concordance (cf. Siegel, 1956, p. 286) do not contain values for seven or fewer rankings of sets of three coefficients, consideration of the values in Table 12 suggests that both W_s would be significant at $p < .01$ had values for these conditions been provided. In sum, the data in Table 12 suggest a considerable amount of construct validity for the notion of innovativeness as assessed by the crude, single-item measures employed in this investigation.

Conclusions

Despite the fact that the range of values was probably restricted by the nature of the subjects employed (i.e., all were young, relatively literate, members of the "Greek" community on the same midwestern campus), the data collected in this investigation argue for acceptance of the following conclusions:

1. Simultaneously applying the self-designating, sociometric, and key informant methods of measuring opinion leadership to relatively cohesive groups reveals that there is a substantial degree of construct validity for the notion of opinion leadership.
2. Opinion leadership tends to overlap different product categories, and the more similar the product categories, the greater the degree of overlap. These findings are consistent with the results of other mono-method investigations of opinion leadership (e.g., King and Summers, 1970; Marcus and Bauer, 1964; Myers and Robertson, 1972; Montgomery and Silk, 1971).
3. Opinion leadership is positively correlated with innovativeness. Again, this is consistent with the results of other investigations (e.g., Myers and Robertson, 1972; Robertson, 1970, pp. 136-137; Rogers and Stanfield, 1968; Tigert and Arnold, 1971).
4. Innovativeness, which can be assessed sociometrically and via key informants, would also appear to possess a high degree of construct validity.

Table 12

Multimethod-Multitrait Matrices for Groups 3 and 4 on Innovativeness

Group 3

	<u>Sociometric</u>			<u>Key Informant</u>	
	Clothing	Cosmetics	Room Dec's	Clothing	Cosmetics
<u>Sociometric</u>					
Cosmetics	.587				
Room Dec's	.266	.236			
<u>Key Informant</u>					
Clothing	.846	.466	.158		
Cosmetics	.590	.683	.665	.496	
Room Dec's	-.073	.070	.665	-.067	.032

Group 4

	<u>Sociometric</u>			<u>Key Informant</u>	
	Clothing	Cosmetics	Room Dec's	Clothing	Cosmetics
<u>Sociometric</u>					
Cosmetics	.510				
Room Dec's	.154	.213			
<u>Key Informant</u>					
Clothing	.509	.666	.067		
Cosmetics	.388	.851	.501	.482	
Room Dec's	-.061	-.019	.553	-.067	.061

Footnotes

1. The assistance of David Giffin and Sue Speaks in collecting these data is appreciatively acknowledged.

References

- Campbell, D. T., & Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105.
- Katz, E., & Lazarsfeld, P. F. Personal influence: The part played by people in the flow of mass communications. Glencoe, Illinois: The Free Press, 1955.
- King, C. W., & Ryan, G. E. Identifying the innovator as a consumer change agent. In D. Gardner (Ed.), Proceedings, 2nd Annual Conference, Association for Consumer Research, 1971, 446-451.
- King, C. W., & Summers, J. O. Overlap of opinion leadership across consumer product categories. Journal of Marketing Research, 1970, 7, 43-50.
- Lazarsfeld, P. F., Berelson, B., & Gaudet, H. The people's choice. 2nd ed. New York: Columbia University Press, 1948.
- Marcus, A. S., & Bauer, R. A. Yes: There are generalized opinion leaders. Public Opinion Quarterly, 1964, 28, 628-632.
- Myers, J. H., & Robertson, T. S. Dimensions of opinion leadership. Journal of Marketing Research, 1972, 9, 41-46.
- Montgomery, D. B., & Silk, A. J. Clusters of consumer interests and opinion leader spheres of influence. Journal of Marketing Research, 1971, 8, 317-321.
- Robertson, T. S. Consumer behavior. Glenview, Illinois: Scott Foresman and Company, 1970.
- Rogers, E. M., & Catarno, D. G. Methods of measuring opinion leadership. Public Opinion Quarterly, 1962, 26, 435-441.
- Rogers, E. M., & Stanfield, J. E. Adoption and diffusion of new products: Emerging generalizations and hypotheses. In F. M. Bass, C. W. King, & E. A. Pessemier (Eds.), Application of the sciences to marketing management. New York: Wiley, 1968, 227-250.
- Siegel, S. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956.
- Tigert, D. J., & Arnold, S. J. Profiling self-designated opinion leaders and self-designated innovators through life style research. In D. Gardner (Ed.), Proceedings, 2nd Annual Conference, Association for Consumer Research, 1971, 425-445.

Appendix A

Self-Designating Opinion Leadership Instrument for Groups 1 and 2

1. During the past school year have you spoken to anyone about new clothing?

- _____ no one
- _____ 1 person
- _____ 2 people
- _____ 3 people
- _____ 4 people
- _____ 5 or more people

2. Compared with your circle of friends, are you more likely, about as likely, or less likely to be asked for advice about new clothing?
- more likely
 about as likely
 less likely
3. Thinking back to your last discussion about new clothing, were you asked for your opinion of the clothing or did you ask someone else?
- I was asked
 I asked someone else
 both
4. When you and your friends discuss new clothing, what part do you play?
- mainly listen
 try to convince them of my ideas
5. Which of these happens more often?
- you tell your friends about new clothing
 your friends tell you about new clothing
6. Do you have the feeling that you are generally regarded by your friends as a good source of advice about new clothing?
- good source
 better than average source
 average source
 below average source
 not a good source at all

Self Designating Opinion Leadership Instrument for Groups 3 and 4

1. In general, do you like to talk about clothes with your friends?
- very much
 to some extent
 very little
 not at all
2. Would you say you give "very little information", "an average amount of information", or "a great deal of information" about clothes to your friends?
- a great deal
 an average amount
 very little
3. During the past school year, have you spoken to anyone about any clothes?
- no one
 1 person
 2 people
 3 people

- 4 people
 5 or more people

4. Compared with your friends, are you "less likely", "about as likely", or "more likely" to be asked for advice about clothes?

- more likely
 about as likely
 less likely

5. If you and your friends were to discuss clothes, what part would you be most likely to play? That is, would you "mainly listen to your friends' ideas" or would you "try to convince them" of your ideas?

- you mainly listen to your friends' ideas
 you try to convince them of your ideas

6. Which of these happens more often: Do you tell your friends about clothes, or do they tell you about clothes?

- you tell them about clothes
 they tell you about clothes

7. As a source of advice about clothes, how do you feel that you are regarded by your friends?

- good source
 better than average source
 average source
 below average source

DOMINANCE AND CONFLICT IN FAMILY
PURCHASING DECISIONS¹

Arch G. Woodside²
University of South Carolina

Marital decision making processes have been studied from four points of view: (1) bases for role differentiation, (2) power structure, (3) decision making structure, and (4) demographic and psychosocial influence on the power structure. Herbst (1952) has suggested four bases for role differentiation: household duties, child control and care, social activities, and economic activities. Parsons and Bales (1955) distinguish between instrumental and expressive roles and, among economic activities, Ferber (1955) draws a dichotomy between "policy" and "routine household" decisions.

Such bases for marital role differentiation have been used to theoretically develop, or empirically explain, categories of marital power structures. The study of power structure has focused on the question of husband or wife dominance. Herbst (1954) developed four decision making power structures: (1) autonomic, or when an equal number of decisions is made by each spouse, (2) husband dominant, (3) wife dominant, and (4) syncratic, or when most decisions are made by both husband and wife. In one empirical investigation Wolgast (1958) concludes that the husband dominates the wife in the decision making process for automobile purchases. She found the wife dominating the husband for household goods and furniture purchasing decisions. Almost perfect agreement in husbands' and wives' reports about relative influence was found by Wolgast.

Davis (1970) demonstrated that the decision structure in the purchase of automobiles is not related to the decision structure in the purchase of furniture. Within each of these product categories, product selection (model, make, color) and allocation decisions (how much to spend and when to buy) were shown to be the structure of the decision process.

Significant demographic and psychosocial relationships with the family power structure have been found to exist. For example, the degree of joint decision making typically declines over a family's life cycle (Wolgast, 1958). Also, when neither husband nor wife belongs to a connected social network they have a greater tendency to engage in joint decision making (Bott, 1957).

Conflict in the marital decision making process has been a poorly studied concept. The comparison of husband dominant, wife dominant, syncratic and autonomic, and conflict family power structures by demographic and psychosocial influence may offer insight into family role differentiation.

Engel, Kollat, and Blackwell (1968) have indicated that the extent of husband-wife involvement varies considerably from product to product. These authors report husbands having a greater tendency to be involved in problem recognition when the product is technically or mechanically complex, as in the case of automobiles, refrigerators, and paint.

The present article focuses on the power structure for the marital decision making process in purchasing eight consumer products: Does dominance in the marital decision making process significantly vary across product categories

being purchased? Does the wife or husband tend to dominate the decision making process for particular products? What products tend to be most closely associated with one another by marital power structures? Answers to these questions should provide a number of theoretical and managerial implications to marketing.

Method

A cross sectional survey of 200 upper-lower, lower-middle, and upper-middle social class families in the Columbia, South Carolina, metropolitan area was taken to study the relationships between role differentiation, dominance, and demographic-psychosocial variables. A random sample of street tracts of single family dwellings was drawn with two families surveyed on each street.

Husbands and wives were interviewed separately in the 200 families. Two interviewers were used, male and female graduate students. One interviewer went into each house separately.

The survey instrument consisted of two parts. Couples answered a series of questions on the relative influence of each spouse for eight products: automobiles, lawnmowers, automatic washing machines, beer, rugs/carpets, cheese, television sets, and gardening supplies. Secondly, the couples completed a demographic and life style instrument which included questions on family life cycle, occupation of family head, wife employment, education, income neighbor visits, popularity, conservatism, club activities, advertising attitudes, opinion leadership, and other life style questions. The life style questions were selected from 300 factor analyzed questions developed by researchers at Purdue University (Tigert, 1969). Typically, four questions were included for each factor.

The decision making questions for the eight products included questions on who first brought up the idea of purchasing, discussion of the purchase with friends, neighbors, relatives, obtaining information from mass media, obtaining information from stores (dealers), style or type, visiting stores or dealer showrooms, specific retail outlet, actual purchase, and experiencing dissatisfaction.

Each spouse was given three answers to choose the relative influence in the decision: husband, wife, husband and wife.

Findings: Examples of Relative Influence

Table 1 is a list of the responses of husbands and wives for two of the eight products in percent. Both the husbands and wives report considerable variation in their roles within the decision making process. For example, 65.5% of the 200 husbands report the wife alone brought up the idea to purchase their automatic washing machines, while only 18.5% report the wife alone made the actual purchase. Differences seem to exist between the two products in spouses' relative influence. Bringing-up the idea to purchase automatic washing machines appears to be a wife-dominant activity, while husbands dominate for this question for television sets. The percentages shown for husbands and wives in Table 1 are similar in size. Not much disagreement is shown between the sexes for the data when grouped. The majority of both the husbands and wives reported the wife alone brought up the idea to purchase automatic washing machines and the husband alone brought up the idea to purchase television sets.

Table 1
 Perceived Marital Roles of Husbands and Wives in Purchasing Automatic Washing Machines
 and Television Sets in Percent (N = 200)

Decision	Automatic Washing Machine						Television Set					
	Husband Responses			Wife Responses			Husband Responses			Wife Responses		
	H	W	J	H	W	J	H	W	J	H	W	J
Brought Up Idea	12	65.5	17.5	11.5	66.5	18.5	52.5	32	14	52	31	16
Style	17	48.5	30	16	51	29.5	45	38.5	14.5	41.5	41.5	16
Size	15.5	50.5	30	19	46	30	48	42.5	7.5	43.5	45.5	10
Brand	22	39	35	19.5	36.5	39.5	51.5	39.5	7	44	45	9
How Much	30	24	42	28	23	44.5	45.5	44.5	8	42.5	45.5	10
Consulted Friends	3.5	10.5	12	3.5	17	13	19	18.5	1.5	10	19	3
Consulted Mass Media	7	10	17.5	3	13	17.5	18	26	3	9.5	24	3.5
Consulted Stores	13	16	36	13	17.5	39.5	31.5	29.5	4	25.5	28.5	5
Visited Stores	14	17.5	60.5	11.5	20.5	62	24	63	10	26.5	58	11.5
Specific Outlet	30	20	45.5	25	25	44	47	43.5	7.5	42	47	9.5
Actual Purchase	39	18.5	38	35.5	23	32.5	44.5	45	8.5	49.5	39.5	10
Experienced Dissatisfaction	0.5	3	0	0.5	3.5	0	1	7.5	1.5	1	9	1

H = Husband
 W = Wife
 J = Joint

Information contained in Table 1 does not answer the question on dominance or conflict and was developed to show an example of relative influence across all families. The study of marital dominance and conflict requires some measures of these variables for each family for the product category or decision area being analyzed.

Indexes of Dominance and Conflict

An index of dominance (I_D) was developed for this study based on the husband and wife responses for the 12 decisions as discussed and listed in Table 1:

$$I_D = \frac{\Sigma(H_a + W_a) - \Sigma(H_b + W_b)}{\Sigma(H_a + W_a + H_b + W_b + H_c + W_c)}$$

where H = Husband's response
 W = Wife's response
 a = Husband alone
 b = Wife alone
 c = Joint

The 12 specific decisions were considered to be of equal importance and no weighting procedure was used to derive a family's I_D score. Assume both the husband and wife reported the husband to be dominant throughout the decision process for buying a television set, the I_D score for this family's decision would be equal to $\frac{(12 + 12) - (0 + 0)}{24} = 1$. If both reported the wife dominant the I_D would equal $\frac{(0 + 0) - (12 + 12)}{24} = -1$.

An index of conflict (I_C) was also developed:

$$I_C = \frac{\Sigma|H_a - W_a| + \Sigma|H_b - W_b| + \Sigma|H_c - W_c|}{\Sigma(H_a + W_a + H_b + W_b + H_c + W_c)}$$

If the husband and wife completely disagree, the I_C would be equal to $\frac{|12 - 0| + |0 - 12| + |0 - 0|}{24} = 1$. If there was complete agreement,

$$I_C = \frac{|12 - 12| + |0 - 0| + |0 - 0|}{24} = 0, \text{ for example.}$$

The I_D or I_C was not calculated for a family for a specific product when the husband and wife had an unequal number of responses or when they answered less than 6 of the 12 responses.

Differences of I_C and I_D means between products were analyzed. Role differentiation theory would suggest significant husband or wife dominance depending on the products' relation to household duties, children, social and economic activities.

Findings: Dominance and Conflict

Table 2 is a list of the average dominance values for the eight product categories. All eight means were statistically significant from zero in testing the null hypothesis that neither the husband nor the wife dominated in the decision making process. Wives dominated the decision making processes for three products: cheese, rugs, and automatic washing machines. Husbands were dominant in the decision making process for television sets, gardening supplies, automobiles, beer, and lawnmowers. The data suggest that products can be meaningfully differentiated by the amount of dominance displayed by either the husband or wife.

Table 2

Mean Dominance Scores (\bar{X}), Standard Deviations (s),
Sample Size (n) and Test of Significance
For $I_D = 0$ For Eight
Product Categories

Product Category	\bar{X}	s	n	t*
Cheese	-.594	.540	187	15.05
Rugs	-.277	.457	179	8.11
Automatic Washing Machines	-.139	.489	198	4.01
Television Sets	.349	.456	199	10.81
Gardening Supplies	.433	.657	188	9.04
Automobile	.534	.366	188	29.25
Beer	.711	.493	108	14.99
Lawnmowers	.781	.359	180	29.25

* t values listed are significant at .001 level.

The dominance averages ranged considerably within -1 to +1. Purchase decisions for lawnmowers and beer tend to be highly dominant by the husband, while automobiles, gardening supplies, and television sets appear to be more moderately dominated by the husband. Wives tend to dominate the purchase decisions for cheese more than they do for rugs or automatic washing machines.

The results appear quite consistent with previous research findings and hypothesized expectations, lending credence to the construct validity of the measurability of the instrument and the methodology:

1. Products to be used outside of the house, e.g., lawnmowers and gardening supplies, tend to require manual work more expected of the husband than the wife.

2. Mechanically complex and expensive product purchases, e.g., automobiles and television sets, are usually made with greater husband, compared to wife involvement (Engel, *et al.*, 1968).

3. Products to be used inside the house and where decor is a concern, e.g., rugs, may tend to reflect the tastes of the wife more than the husband.

4. Assuming the wife's dominance in "doing the wash" even though some of the husbands did appear to wear the apron strings, the wife would be expected to have the final words in the purchase of an automatic washing machine.

Table 3 is a symmetrical matrix of product moment correlation coefficients of every product with every other product.

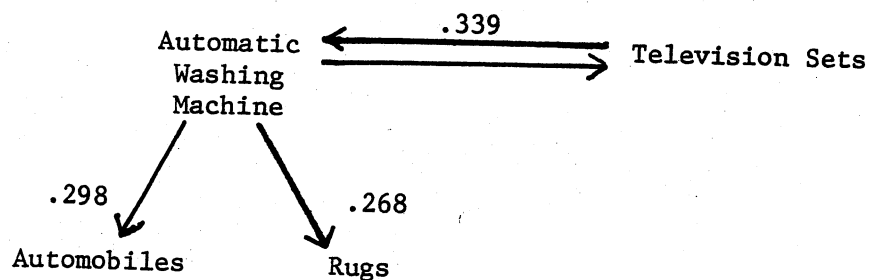
A number of correlation coefficients are negative in Table 3. This would suggest that products affect choice of answers by index of dominance values. Groups of families who consistently rate the husband or wife dominant do not appear to be present. Products do affect choice of answers by respondents.

Product clusters were formed following the procedure outlined by Kamen (1970) from the correlations in Table 3. Figure I is a description of the two clusters that resulted with dominance being the dependent variable.

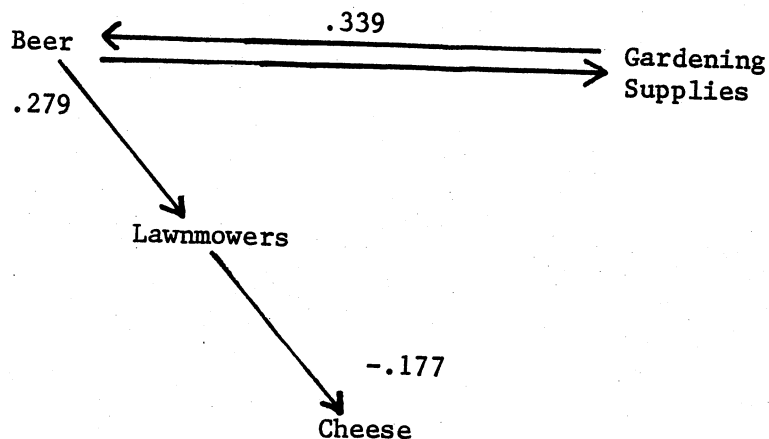
Figure I

Clustering of Products by Marital Dominance

Cluster 1:



Cluster 11:



Cluster 1 consists of four durable products, each representing sizeable income expenditures for the family. The correlation coefficients in Cluster 1 are statistically significant ($p < .05$) and suggest similar dominance tendencies across the four product categories.

Table 3
Correlations Among Product Dominance Scores (n = 80)

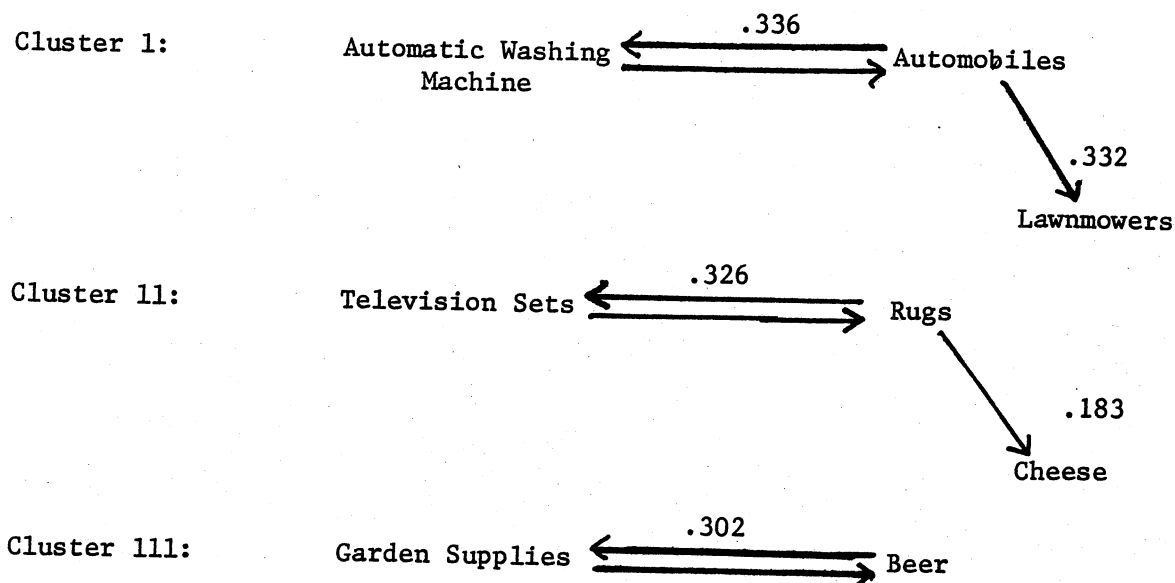
Product	Automatic Washing Machine	Television Sets	Automobiles	Lawnmowers	Beer	Rugs	Cheese	Garden Supplies
Automatic Washing Machine	-	.339	.298	.082	.020	.268	.075	.065
Television Sets	.339	-	.273	.125	-.008	.211	.010	-.002
Automobiles	.298	.273	-	.080	-.129	-.030	.075	-.232
Lawnmowers	.082	.125	.080	-	.279	-.010	-.177	.248
Beer	.020	-.008	-.129	.279	-	.050	.033	.339
Rugs	.268	.211	-.030	-.010	.050	-	-.097	.142
Cheese	.075	.010	-.075	-.177	.033	-.097	-	.004
Garden Supplies	.065	-.002	-.232	.248	.339	.142	-.004	-

Cluster 11 consists of four semi-durable and nondurable products representing lower priced items than those appearing in Cluster 1. The correlation coefficients in Cluster 11 are statistically significant ($p < .05$) except for lawnmowers and cheese ($r = -.177, p > .05$). Cluster 11 appears to be a catch-all grouping of products defined, in part, by the limited number of products examined.

Table 4 includes correlation coefficients of index of conflict scores for the families used to compute Table 3. Mean index of conflict values and standard deviations are also shown in Table 4. Figure II consists of clusters formed from the conflict correlation coefficients in Table 4.

Figure II

Clustering of Products by Marital Conflict



The correlation coefficients shown in Figure II are statistically significant ($p < .05$) except for rugs and cheese ($r = .183, p > .05$).

The three clusters formed represent a different configuration for the products where conflict is the dependent variable compared to Figure I where dominance is the dependent variable. This suggests that power structures among products change according to the variable of power being analyzed.

Cluster 1 in Figure II consists of three products mechanically complex and durable. Automobiles and automatic washing machines are found in Cluster 1 of Figure II and Cluster 1 of Figure I which indicates a highly similar power relationship between these products.

Cluster 11 in Figure II contains two home furnishing products which are statistically significantly related ($r = .326, p < .01$). Cluster 111 may represent another catch-all grouping of products; however, the relationship shown for gardening supplies and beer was statistically significant ($r = .302, p < .01$).

Table 4

Correlations Among Product Conflict Scores (n=80) and Means (\bar{X}), Standard Deviations (s), and Sample Sizes (n)

Product	Automatic Washing Machine	Television Sets	Automobiles	Lawn-mowers	Beer	Rugs	Cheese	Garden Supplies
Automatic Washing Machine	-	.129	.336	.144	.176	.213	.061	.188
Television Sets	.129	-	.092	.183	-.036	.325	.173	.118
Automobiles	.336	.092	-	.332	.119	.023	-.005	-.092
Lawnmowers	.144	.183	.332	-	.200	.112	-.069	.014
Beer	.176	-.036	.119	.200	-	-.035	.166	.302
Rugs	.213	.326	.023	.112	-.035	-	.183	.215
Cheese	.061	.173	-.005	-.069	.166	.183	-	.133
Garden Supplies	.188	.118	-.092	.014	.302	.215	.133	-
\bar{X}	.266	.279	.207	.172	.180	.231	.229	.257
s	.241	.266	.234	.226	.249	.225	.284	.283
n	194	194	194	187	112	196	192	196

Conclusions

The analyses of dominance and conflict values indicate husband or wife dominance in the decision-making processes of a number of consumer products. The level of husband or wife dominance appears to be a matter of degree, varying by type of product.

Several products may be conceptualized under more general categories and therefore product types can be generalized by family power structures.

Footnotes

1. Data collection assistance of Patty Richardson and Timothy Bice is gratefully acknowledged.
2. Assistant Professor of Business Administration.

References

- Bott, Elizabeth. Family and social network. London: Tavistock Publication, Ltd., 1957.
- David, Harry L. Dimensions of Marital Roles in Consumer Decision Making. Journal of Marketing Research, 1970, 7, 168-177.
- Engel, James F., David T. Kollat, & Roger D. Blackwell. Consumer behavior. New York: Holt, Rinehart and Winston, Inc., 1968.
- Ferber, Robert. On the Reliability of Purchase Influence Studies. Journal of Marketing, 1955, 19, 225-232.
- Herbst, P. G. The Measurement of Family Relationships. Human Relations, 1952, 5, 3-35.
- Herbst, P. G. Conceptual Framework for Studying the Family. Social structure and personality in a city. London: Routledge and Kegan Panel, 1954.
- Kamen, Joseph M. Quick clustering. Journal of Marketing Research, 1970, 7, 199-204.
- Parsons, Talcott, & Robert F. Bales, Family, socialization and interaction process. Glencoe, Illinois: The Free Press, 1955.
- Tigert, Douglas. A Psychographic Profile of Magazine Audiences: An Investigation of a Media's Climate. Unpublished paper presented to Consumer Behavior Workshop. Columbus, Ohio: Ohio State University, 1969.
- Wolgast, Elizabeth H. Do Husbands or Wives Make the Purchasing Decisions? Journal of Marketing, 1958, 23, 151-158.

WHY THE POOR MAY PAY MORE FOR FOOD:
THEORETICAL AND EMPIRICAL EVIDENCE

Howard Kunreuther
University of Pennsylvania

Introduction

Do the poor pay higher prices for food than their wealthier neighbors? A number of empirical studies have been undertaken in recent years on food price differentials across stores in metropolitan areas and have shed considerable light on this questions (Sexton, 1971). To date the most detailed statistical analysis on whether the urban poor pay more for food than do urban residents with higher incomes is the Alcaly and Klevorick (1971) study of food prices in New York City. On the basis of their statistical regression results for 31 commodities, they conclude that the price of a given commodity in a particular type of store tends to be unaffected by or even rises with an increase in the level of neighborhood income. They also point out, however, that the mean price of each of the items studied is higher in the small independent stores than in the chain stores. To the extent that poor people shop at these smaller stores, they will pay higher prices for the same quality food than if they had purchased their groceries at a chain store.

The Alcaly and Klevorick study has raised a number of interesting questions which deserve further theoretical and empirical study. This paper is an attempt to extend their analysis in both these directions. On the theoretical side, I will suggest a framework for better understanding consumer purchasing decisions by considering grocery items which are packaged in several different sizes. At the empirical level, I will report on the results of a study of household food purchasing decisions undertaken in New Haven during the summer of 1971 which was designed to test these theoretical ideas.

Store and Size Effects

By standardizing for quality of food items, we should be able to quantify some of the relevant differences in the purchasing behavior of low- and middle-income shoppers. For any given brand, two principal factors to be considered are the "store effect" and the "size effect." The store effect refers to price differentials between stores for the same-sized item. If the price per ounce for any given package size varies inversely with the size of store, then individuals who shop in chains would pay less for identical items than those who patronize smaller grocers. The size effect refers to differences in price per ounce for various sizes of a particular branded item within any given store. If price per ounce varies inversely with size, then poor people who purchase small packages will pay more for the same quantity in the long run than if they had brought a larger size.

It is important to distinguish between these two factors in analyzing purchasing decisions. The store effect provides a measure of the importance of a consumer's location and his mobility on purchasing decisions. We will want to determine not only what price differentials exist between large and small stores but also whether certain offsetting benefits are provided to patrons of a small neighborhood grocer. For example, if shoppers are permitted to charge their purchases then this could partially account for the higher prices in these stores.

The size effect measures the role which budget, storage constraints and costs of holding inventory play in purchasing decisions. If low-income families buy a market basket of goods on their weekly or bi-weekly shopping trip, then they may be forced to purchase smaller sizes as a consequence of a budget constraint. Restrictions imposed by the size of a shopping cart used to carry purchases from the store as well as by limited food storage space in the home will also operate in the same direction. Low consumption rates will discourage large size purchases due to the cost of holding inventory.

To measure the relative importance of store and size effects on purchasing decisions, eight items packaged in at least three different sizes were chosen for detailed study in New Haven. Their prices and availability were tabulated for 11 chain and 11 local neighborhood stores during the week of July 12-16, 1971 (See Appendix A for store descriptions). Table 1 presents the average price per ounce for each of the sizes as well as the differences between the large and small stores. For all items surveyed, the average price per ounce declined or remained the same whenever package size increased. Although there were occasional exceptions to this general rule within individual stores, these data indicate the presence of quantity discounts. Looking across all stores surveyed, the percentage savings between purchasing the largest instead of the smallest size ranged from 15 per cent (detergents) to over 50 per cent (mayonnaise), indicating that the size effect may be quite pronounced for packaged goods.

For any given size, there is also a store effect for each of the items, as seen by comparing average price data between large stores and small neighborhood stores. These data are consistent with all of the recent surveys on food-price differentials (Bureau of Labor Statistics, 1966; and Federal Trade Commission, 1969). Table 1 also shows that the smaller neighborhood grocers stock fewer larger-sized items relative to the chain stores. This is undoubtedly due to a combination of their own space constraints and their shoppers' preferences for smaller sizes. Individuals shopping in smaller stores thus pay higher prices than the chain shoppers for identical items, while also having a narrower range of choice.²

A Model of Household Food Purchasing Decisions

The data from the New Haven stores suggests that there are significant store and size effects which should affect the family's food purchasing decisions. A simple model incorporating both these factors will now be developed using concepts from the theory of consumer demand.

Consider a family who has a choice of purchasing a combination of n different food commodities. Defining q_j to be the consumption rate of item j per unit of time the family's function for the n commodities is

$$U(q_1, \dots, q_n) . \quad (1)$$

We will make the simplifying assumption, which from our New Haven consumer survey results does not appear unrealistic, that the household has a fixed food budget (B) per unit of time. If p_j is the price for item j , the budget constraint can be written as

$$\sum_{j=1}^n p_j q_j = B . \quad (2)$$

Table 1

Grocery Store Prices in New Haven Stores (July, 1971) For 11 Chain
And Large Independents and 11 Small Neighborhood Stores

Item	Size (in ounces)	Chain and Large Independents			Small Neighborhood Stores		
		Average Price Per Ounce (All Stores) (in cents)	Number of Stores Stocking Item	Average Price Per Ounce (in cents)	Number of Stores Stocking Item	Average Price Per Ounce (in cents)	
Mott's Apple- sauce	8	2.37	10	2.20	5	2.54	
	15	1.77	11	1.63	9	1.86	
	25	1.69	8	1.51	5	1.84	
	35	1.45	11	1.35	7	1.54	
	48	1.23	1	1.23	0	--	
Heinz Catsup	12	2.67	7	2.42	5	2.87	
	14	2.30	10	1.97	7	2.59	
	20	2.09	11	2.01	5	2.19	
	26	1.90	9	1.75	2	2.27	
	32	1.79	9	1.78	1	1.84	
	48	1.44	9	4.19	8	4.57	
Hellman's Mayonnaise	16	3.17	10	2.92	8	3.36	
	32	2.54	11	2.28	5	2.77	
	48	2.54	5	2.38	3	2.64	
	128	2.15	1	2.15	0	--	
	6	5.29	6	4.94	5	5.50	
Skippy Peanut Butter	12	4.25	11	3.88	9	4.49	
	18	3.85	11	3.56	4	4.28	
	28	3.63	11	3.37	4	4.03	
	40	3.58	6	3.51	2	3.68	
Maxwell House Instant Coffee	2	29.78	9	28.85	5	30.70	
	6	20.41	9	18.57	6	21.95	
	10	17.99	11	16.09	5	20.30	
Kellogg's Corn Flakes	8	3.27	10	2.98	6	3.56	
	12	2.98	8	2.85	7	3.08	
	18	2.41	11	2.41	3	2.41	
Pillsbury Flour	32	1.09	9	1.05	6	1.12	
	80	.83	11	.77	8	.87	
	160	.79	9	.71	2	.96	
	400	.62	5	.62	3	.63	
Tide Detergent	20	2.08	8	1.95	5	2.21	
	49	1.89	9	1.81	5	1.96	
	84	1.79	9	1.72	2	1.98	
	171	1.75	6	1.65	1	1.92	

We can derive a demand curve for item j by holding all prices constant except p_j and maximizing (1) subject to the budget constraint of (2).³

The size effect on consumer purchasing decisions can be incorporated into the analysis by deriving an implicit "supply" schedule for each good. To see this more clearly, assume that the j^{th} item is packaged in m different sizes, 1 being the smallest and m the largest. Let size i contain k_j^i ounces and sell at a price per ounce p_j^i . If a quantity discount phenomenon exists, then $p_j^{i+1} < p_j^i$. Since the package size i lasts for k_j^i/q_j time units, there is an inventory cost associated with storing the unused portion of the good. Let h represent the percentage cost per dollar per unit of time.⁴ The cost per ounce of the i^{th} size per unit of time is then given by

$$C_j^i = p_j^i + \frac{hp_j^i}{k_j^i} \int_0^{k_j^i/q_j} \left[\frac{k_j^i}{q_j} - t \right] q_j dt = p_j^i + \frac{hp_j^i k_j^i}{2q_j} \quad (3)$$

$$i = 1, \dots, m$$

$$j = 1, \dots, n$$

For each item j the household will want to choose the size i which minimizes C_j^i . From (3) we see that if a quantity discount phenomenon exists, then optimal package size increases as the consumption rate (q_j) increases. For a given value of h we can trace out a supply schedule for item j which shows the lowest economic cost per ounce as a function of q_j . For specificity, let the unit of time be one year and $h = .1$. The solid line in Figure 1 represents the supply curve, SS, for catsup based on price data from a chain store in the New Haven area which stocked five different sizes. The points of discontinuity indicate that another package size has become optimal, based on (3).⁵ The choice of package size for item j is determined by the intersection of the consumer's demand schedule for catsup, DD, with SS as illustrated by point E_1 in Figure 1. Such an intersection will represent a stable equilibrium if the demand schedule cuts the supply schedule from above.

Several implications of the size effect on consumer purchasing decisions can be derived from this model. From (1) and (2) it can easily be shown that a decrease in B will cause a leftward shift in the demand curves for goods which have a positive income elasticity of demand. If quantity discounts exist, then the optimal purchase size of these items will vary directly with B . A constraint on the amount of space for food can be incorporated through h , which will increase to reflect both the opportunity cost of money as well as a marginal value of extra storage space. From (3) we see that an increase in h will shift the supply curve upward and yield the same or smaller optimal package size for any given value of q_j . If low-income families have the most severe storage and budget restrictions, then the above analysis would predict that for any given q_j they will purchase the same or smaller packages than middle-income families.

The store effect is reflected in different values of p_j^i across stores and will lead to a shift in the supply curve. To illustrate, consider the dashed supply curve, S'S', in Figure 1 based on prices of catsup in a small local store in New Haven which only stocked three sizes. The price per ounce is higher for each size relative to the chain store, providing a clear

\$ per year

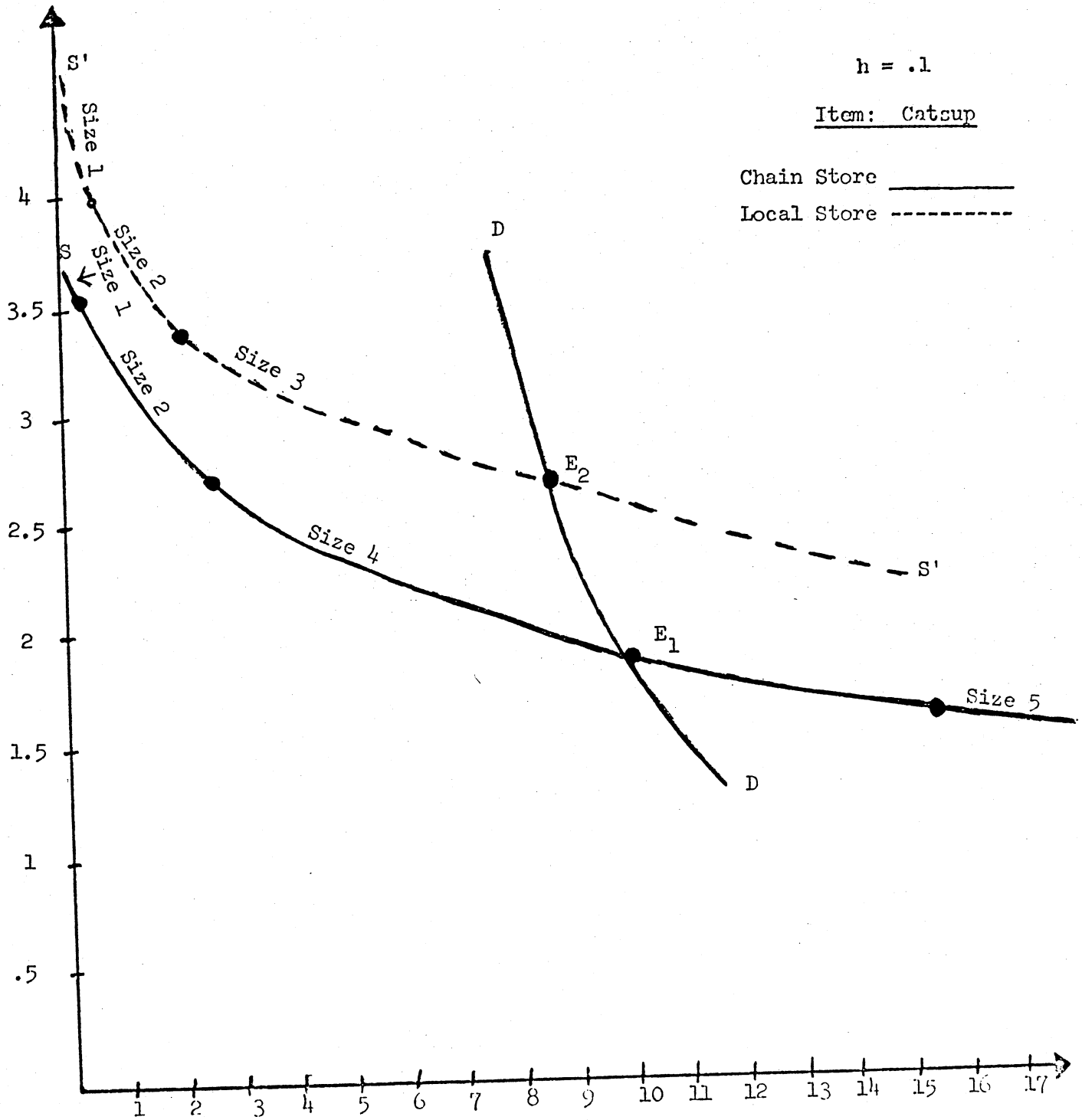


Fig. 1.--Optimal purchase decision as a function of size and store effect.

illustration of the store effect. In this case, the equilibrium point for the demand curve given by DD is at E_2 indicating that size 3, rather than size 4, would have been the optimal one to purchase.

What factors determine the choice of store at which a consumer will purchase his goods? The costs associated with travelling to and from a store may play an important role in the shopper's selection process. Looking at a specific trip, low-income families who have limited mobility will have much higher transportation costs per mile for long-distance travel than middle-income families with automobiles. They are thus more likely to shop at the neighborhood store than to travel some distance to chain stores. In formal terms, suppose a consumer has the choice of shopping at r different food stores. If F_k is the transportation cost per unit of time associated with travelling to and from store k then (2) becomes

$$\sum_{j=1}^n P_j q_j - F_k = B, \quad k = 1, \dots, r. \quad (2')$$

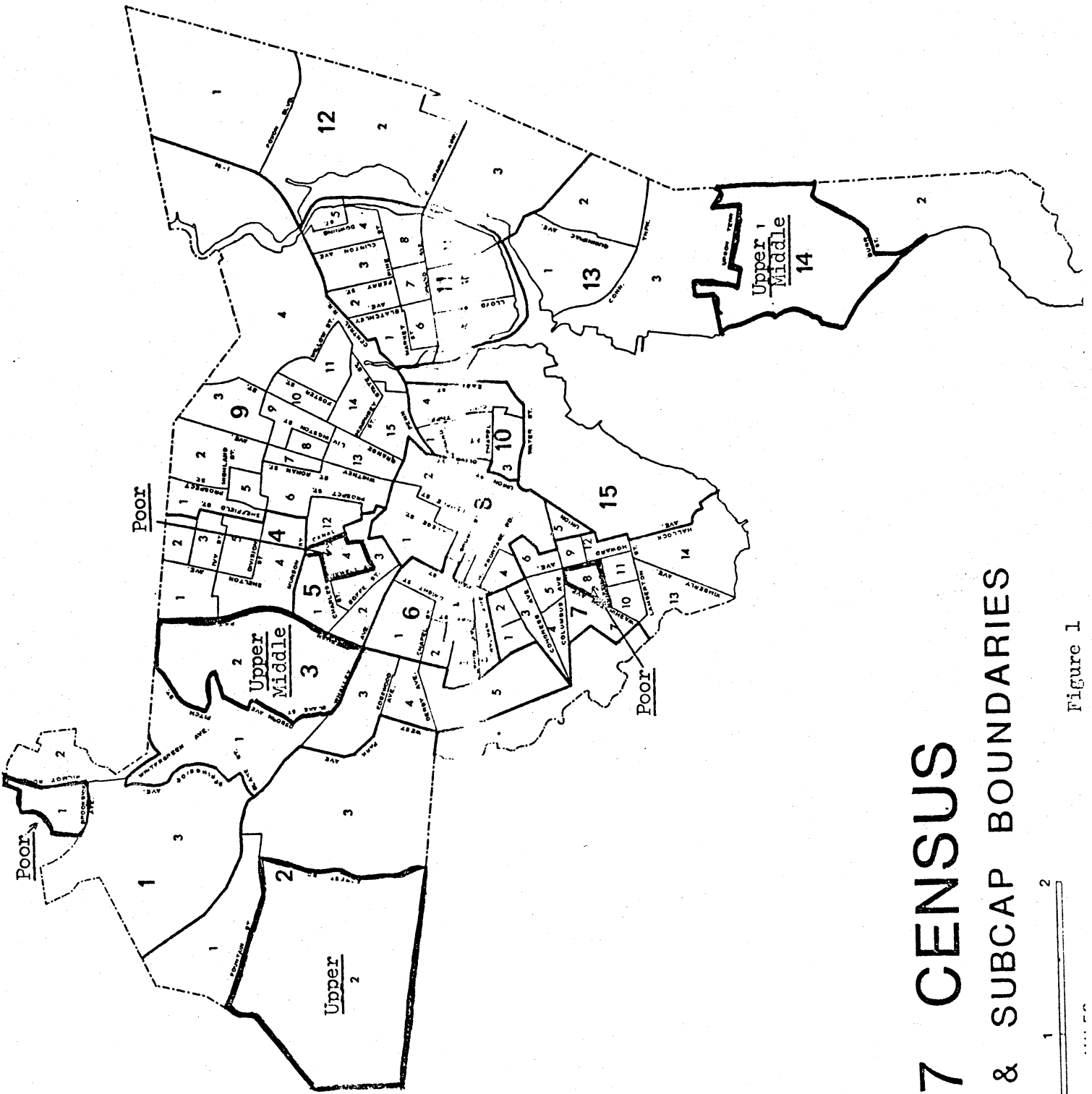
To illustrate how (2') may operate, consider a simple example where $r = 2$, with store 1 being the nearest chain store and store 2 the local grocer (assumed to be much closer to the household than store 1). If F_1 is sufficiently greater than F_2 , it would be optimal for the poor family to shop at store 2 since the decrease in transportation cost would more than offset the higher food prices at the local store.

Empirical Tests of the Model

To better understand consumers' purchasing decisions and test these theoretical ideas we undertook a survey of 159 households in the New Haven area. A 25% random sample of New Haven residents had been undertaken in 1967 to pretest the 1970 U. S. Census, and for this purpose the city was divided into 13 cap districts with each cap then subdivided into a number of subcaps. Figure 2 details these boundaries and the six subcaps we surveyed. Some of the detailed economic data which were collected in 1967 are summarized in Table 2. As can be seen from the income data the first three subcaps (2-2, 3-2, 14-1) are relatively well-to-do areas of the city while the other three (7-8, 1-1, 5-4) are relatively poor sections. By combining these subcaps to form an upper-middle-income group and a low-income group it is possible for us to make preliminary comparisons as to differences in food purchasing patterns as a function of income.

We interviewed between 25 and 30 families chosen on a quota sampling basis in each of the subcaps to better understand their shopping patterns and their purchase size decision (See questionnaire in Appendix B). Except for subcap 2-2 (where a number of residents were reluctant to answer doorbells), approximately 90 per cent of the housewives were willing to be interviewed. A Malaysian student interviewed in the black areas and two white female graduate students covered the other subcaps. The final sample comprised 78 housewives in the middle-income group and 81 in the low-income bracket.

Based on the above consumer purchasing model, two hypotheses, one related to the store effect and the other to the size effect, were formulated:



1967 CENSUS CAP & SUBCAP BOUNDARIES

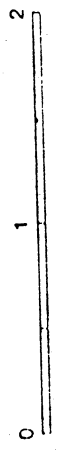


Figure 1

TABLE 2
ECONOMIC DATA ON SIX SUBCAPS IN NEW HAVEN

Group	Subcaps	Average			Housing Data			Population Data			
		Income	Family Size	No. of Renters	Average Rent Per Month	No. of Home-Owners	Average House Value	Total	White	Black	Other
I	2-2	16,548	3.3	241	160	747	33,243	3,106	3,048	29	29
	3-2	10,113	3.2	788	118	816	27,282	4,645	3,868	753	24
	14-1	9,862	3.6	192	148	755	23,265	3,158	3,154	4	0
II	7-8	5,592	4.2	450	105	136	16,743	1,842	1,073	769	0
	1-1	4,460	5.1	164	75	0	0	773	151	622	0
	5-4	3,733	4.6	696	69	25	25,650	2,397	150	2,242	5

Source: 1967 Special Census of New Haven.

Store Effect Hypothesis. Low-income families will be more limited in mobility than middle-income families and hence will be more prone to shop in stores close to their homes. Given the relatively few chain stores in these areas the percentage of low-income consumers shopping at small neighborhood stores will be significantly higher than for middle-income consumers.

Size Effect Hypothesis. Consider a low- and middle-income household each with identical family size and identical per capita consumption rates for a given item. Due to more stringent budget and storage constraints, the low-income family will tend to purchase a smaller-sized package than will a middle-income family. As family size increases we would expect to find a higher total consumption rate per unit of time for most items and hence a preference for larger-sized packages.

Table 3 presents summary data on the store effect for the two income groups. Only five per cent of middle-income families shop at small neighborhood stores while over 60 per cent of the low-income families do their primary shopping at the small local grocer. One reason for this difference in shopping patterns is provided by the data on mode of transportation and distance travelled to the store. A sizable proportion of low-income families use public transportation or walk to their local grocery store while practically all middle-income families drive automobiles to a chain. This is also reflected in figures on the average distance travelled: 1.9 miles for the low-income group and 3.6 miles for the middle-income group.⁶

Due to the limitations of the questionnaire data we were only partially successful in measuring the magnitude of the size effect. Our own budget and time constraints prevented us from making detailed pantry checks so that we had to rely on perceptual answers regarding "number of sizes available in your store," "size of purchase" and "frequency of purchase." Interestingly enough, Table 4 indicates that for the eight items chosen for detailed study the average low-income family perceived at least the same number of available sizes as the middle-income shopper did. Despite the larger average family size in low-income households, their ratio of actual purchase size to number of sizes perceived was lower than for middle-income families on all items except applesauce.

Table 4 also shows that low-income families purchased all eight items more frequently than middle-income families. This could be due to the presence of budget and storage constraints and/or different consumption rates for the two groups. From Table 4 we can see that low-income families had more restrictive per capita food budgets on the average than middle-income families and a greater percentage of them perceived inadequate storage space. In this sense, the evidence from the questionnaire supports the size effect hypothesis and suggests that the poor may pay more because constraints force them to buy smaller sizes on a more frequent basis than middle-income shoppers.

How do the New Haven results compare with other surveys on consumer behavior? To my knowledge the only study which explicitly examines the choice of store by consumers is one undertaken by Goodman (1968) in a low-income area of Philadelphia where there were no large or modern food retailing facilities. In contrast to the New Haven survey, he found that approximately 92 per cent of the 520 families interviewed did their principal grocery shopping outside of their neighborhood and most of these families shopped at supermarkets or medium-sized independents, using the local stores as supplementary sources of food. With respect to mode of transportation, approximately

TABLE 3
 COMPARISON OF STORE EFFECT BETWEEN MIDDLE- AND
 LOW-INCOME GROUPS IN NEW HAVEN

	<u>Middle-Income</u> (Group I)	<u>Low-Income</u> (Group II)
<u>Primary Store for Shopping</u>		
Local	11 (14) ^a	52 (64)
Chain	67 (86)	29 (36)
<u>Mode of Transportation</u>		
Automobile	74 (95)	41 (51)
Foot	4 (5)	34 (42)
Public Transportation	0 (0)	6 (7)
Average Distance Travelled (in miles)	3.6	1.9

^a() = percent of total.

Source: Responses to questionnaire on grocery store purchases (see Appendix B).

TABLE 4

COMPARISON OF SIZE EFFECT BETWEEN LOW- AND MIDDLE-
INCOME GROUPS IN NEW HAVEN

Item	Middle Income (Group I)		Low Income (Group II)		Average Time Between Purchases (in weeks)
	Average No. of Sizes Perceived	Actual Purchase Size	Average No. of sizes Perceived	Purchase Size	
Applesauce	2.9	.76	3.0	.80	2.3
Catsup	2.8	.82	2.9	.79	2.4
Mayonnaise	2.8	.86	3.1	.81	3.1
Peanut Butter	3.1	.81	3.1	.77	3.0
Instant Coffee	2.8	.86	2.9	.76	3.8
Cold Cereal	2.7	.81	3.0	.80	1.2
Flour	2.9	.62	3.2	.61	3.7
Detergent	3.7	.76	3.7	.70	2.5
<u>Adequate Storage Space</u>					
	Yes	68 (87)		33 (41)	
	No	10 (13)		48 (59)	
<u>Food Budget</u>					
Average weekly expenditures		39.4		36.6	
Average family size		3.4		4.6	
Average per capita weekly expenditures		11.5		7.9	
Use of Food Stamps	Yes	2 (2.5)		35 (43)	
	No	76 (97.5)		46 (57)	

45 per cent of the sample used automobiles, an additional 14 per cent used public transportation and the remainder walked to the store. Since over 40 per cent of the residents shopped by foot, these data imply that larger stores were located relatively close to the low-income area, in contrast to New Haven.⁷

Empirical evidence on the size effect has been presented by Frank Douglas and Polli (1967) based on 491 households from the Chicago Tribune 1961 consumer panel survey. Multiple regression techniques were utilized to explain the per cent of purchases made of small package sizes by the ith household for each of 31 different grocery products. Among the 11 variables which they found to be statistically significant, purchase size was positively correlated with income and family size and negatively related to building size. All these relationships are consistent with the predictions from the theoretical model of consumer behavior developed in the previous section.

Suggestions for Future Research

This paper should be viewed as a first effort in developing a formal framework for understanding consumer purchasing decisions. The questionnaire administered in New Haven uncovered a number of other factors which appear to play an important role in the choice of store and package size.

Value of Time. Consumers' value of time may play a role in purchase size decisions. Schraier (1972), building on the work of Becker (1965), has investigated this question. He suggests that shoppers who have a high value of time will tend to purchase large-sized items and make relatively few trips to the store. It is likely that these families will have an automobile and can thus easily transport large packages to their residences. Those who must walk to the store are limited with respect to the amount they can carry, and will thus be less prone to take advantage of quantity discounts. It is likely that these families will have a relatively low income, and hence a low value of time, which also suggests that they would make more frequent trips to the store.

Cost of Search. A closely related factor is the cost of search and differences in behavior between income groups. High-income families may have greater mobility than poor families which facilitates their searching process, but this comparative advantage may be offset by their higher value of time.

Price Comparisons Across Sizes. Many stores have recently posted the price per ounce on packaged goods either voluntarily or due to legislation. Proponents of unit pricing legislation claim that burdensome arithmetic computations prohibit shoppers from making meaningful price comparisons when "price per ounce" figures are not posted. If this is true then many individuals are unaware of the extent of quantity discounts and hence may make uneconomical purchases. It should be possible to test this hypothesis by seeing whether or not there are significant differences in the total sales of each size of an item before and after the change to "unit pricing" by a particular store. These data would provide some measure of the value of this additional pricing information to the consumer.

Brand Effect. What differences exist between income classes with respect to the types of brands they purchase? If low-income families purchase nationally advertised brands because their local stores do not stock any house brands, then they may be paying more per ounce for the same quality product.

Market Basket Effect. We have implicitly assumed that each household allocates a proportion of its income for food and then determines a market basket of goods to purchase with this fixed amount. More research is needed to test the realism of this assumption. If an individual were not constrained by a short-run budget, then presumably he could buy a few large packages each shopping trip rather than a market basket of goods unless limited storage space prevented him from making these purchases.

Consumption Rate Effect. We have also assumed in this analysis that the household's consumption rate did not vary with purchase size. One low-income housewife in New Haven claimed she purchased the medium size box of corn flakes on her weekly shopping trip because she could not afford to have her children finish a large size box each week. It would be interesting to determine whether, in fact, poor families purchase some items in small sizes as a way of reducing their consumption rate and hence meeting a long-run income constraint.

Footnotes

1. I would like to express my appreciation to Palomona Ferris, Annette Steyer, and Jomo Sunderan for their help in interviewing residents in the New Haven area. William Wells provided helpful comments on an earlier draft of this paper.
2. Store owners of the small independents did indicate that they provided credit to their regular shoppers but did not state the explicit terms, perhaps because they had no formal rules of behavior.
3. For a standard treatment of this problem, see J. R. Hicks, Value and Capital (Second Edition, London: Oxford University Press, 1957).
4. There is no reason that h must remain the same for each item. Items which deteriorate more rapidly in quality could have a higher value of h reflecting the spoilage factor. A more detailed discussion of inventory costs as it affects consumer demand appears in H. Kunreuther, "The Effect of Quantity Discounts on Consumer Demand: An Application of Inventory Theory," Center for Mathematical Studies in Business and Economic Report No. 7009, University of Chicago, Graduate School of Business, 1970.
5. Equation (3) indicates that it would never be optimal to purchase size 3 for the set of catsup prices in this particular store.
6. If there were fewer stores per square mile in the middle-income area in comparison with the low-income sections, then this would provide an alternative explanation with respect to differences in distance travelled.
7. A study by Sexton may also yield data on the store effect. He restricted his analysis to a comparison of mean prices of three products purchased in the same type of store (e.g., chain, independent) by approximately 220 black and 600 white families. No analysis was made of choice of store by income group but it should be possible to obtain these figures from his dissertation. See Donald E. Sexton, Jr., "Do Blacks Pay More? A Comparison of Prices Paid for Grocery Store Commodities by Black and White Families," unpublished doctoral dissertation, University of Chicago, Chicago, Illinois, 1970.

Appendix A

New Haven Grocery Stores

<u>Area</u>	<u>Neighborhood</u>	<u>Name</u>	<u>Address</u>	<u>Hours</u>	<u>Type of Store</u>
Downtown	Between poor and high rise, upper middle class	Pegnataro's	82 York St.	Mon-Fri 9-9 Sat 9-6	Chain
Downtown Yale Campus	Student and upper middle class	Quality Grocer	65 Broadway	Tue-Sat 9-6	Local
Downtown	Middle to lower middle class	Co-operative Consumers (Conn. Chain)	732 Chapel	Mon-Tue Tue-Wed 9-6 Thurs- 9-7:30 Sat 9-6	Chain
Downtown	Middle to lower middle class	Mohawk Meat Center	406 State St.	Mon-Wed Fri, Sat 9-6 Thur 9-7	Local
Hamden Mart (Golden Mile)	Middle to upper class	Stop & Shop	2300 Dixwell	Mon-Sat 9-9	Chain
Hamden (Golden Mile)	Middle to upper class	Shop Rite	2165 Dixwell	Mon-Fri 9-9 Sat 8-9	Chain
Hamden (Golden Mile)	Upper-middle to upper class	A & P	2845 Dixwell	Mon-Wed 9-9 Tue-Fri 8:30-9 Sat 8:30-6	Chain
Hamden Mart (Golden Mile)	Lower middle to upper middle class	Pegnataro's	2201 Dixwell	Mon-Fri 9-9 Sat 9-6	Chain
Hamden Acme Mall	Middle class	Pegnataro's	1255 Dixwell	Mon-Sat 9-9	Chain
Hamden	Middle class	Stop & Shop	940 Dixwell	Mon-Sat 9-9	Chain
Dixwell	Black ghetto	Lake Vanity	29 Dixwell	Everyday 6-9	Local
Dixwell	Black ghetto	Neighborhood Grocery	77 Dixwell	Tue-Sun 8-9	Local
Dixwell	Black ghetto	Capitol Market	168 Dixwell	Mon-Wed 8-6 Thur-Fri 8-7:30 Sat 8-6	Local
Hill	Poor Latins Italians, Blacks	Pacelli Bros.	178 Liberty	Wed-Sat 8-6	Local
Hill	Poor Latins Italians, Blacks	Danny's Self Service	284 Putnam	Mon-Thur 8-6 Fri 8-7 Sat 8-6	Local

<u>Area</u>	<u>Neighborhood</u>	<u>Name</u>	<u>Address</u>	<u>Hours</u>	<u>Type of Sto</u>
Orange St.	Middle to upper middle class white	Orange Food Market	721 Orange	Mon-Thur 8-6 Fri 8-9	Local
Whalley Ave.	Upper middle class	Pegnataro's	464 Whalley	9-9 9-6	Chain
Westville	Upper middle to upper class	Winton's Market	923 Elm St.	Mon-Sat 8-7	Local
Westville	Upper middle to upper class	Cumberland Farms	796 Edgewood	Everyday 9-11	Local
Westville	Upper middle to upper class	Westville Quality Mkt.	243 Alden Ave.	Mon-Sat 7:30-6 Fri 7:30-7	Local
Westville Amity	Middle class	Stop & Shop	7154 Amity Rd.	Mon-Fri 9-6 9-9	Chain
Westville Amity	Middle class	First Nat'l Stores	112 Amity Rd.	Mon-Fri 9-9 Sat 9-6	Chain

Appendix B

New Haven Questionnaire on Grocery Store Purchases

1. Do you do the shopping for the family? Yes No

2. Where do you do most of your grocery shopping?

Name Name

Location Location

Small neighborhood store Local supermarket Chain supermarket

Why do you shop primarily at this store?

Convenience Lower prices Better selection

Better quality Stamps Charge account Food stamps

Know the people and owner Other (specify)

Any other reasons?

3. How far is the store from your home?

miles minutes

How do you normally get there?

Car Bus Foot Taxi

4. Do you look in the newspaper for special sales on items in different stores? Yes No Occasionally

Do you clip coupons from the newspapers? Yes No Occasionally

If yes, do you go to a different store than the one you normally shop if it has special sales? Yes No

If yes, Occasionally or Frequently

For the following items what sizes are available and which do you normally buy at the store in which you do most of your shopping?

Item	How often do you buy it?	How many sizes are there in your store?	What size do you buy*	No. of ozs. (if specified by housewife)	Price	Convenience	Variety	Spoilage	Consumption Rate Little	Constraints Budget Storage	Too Heavy to carry	Other Specify
Apple-sauce												
Catsup												
Jams or Jelly												
Mayonnaise												
Peanut butter												
Ground coffee												
Instant coffee												
Cold cereal												
Flour												
Rice												
Pork & beans												
Bleach												
Detergent												
Toothpaste												

* let 1 be the smallest size

N.A. = not available

6. How often do you go shopping? Every two weeks _____ Weekly _____
 Twice a week _____ Three times a week _____ Four times a week _____
7. Do you think large size items are cheaper per ounce than small size ones?
 Yes _____ Frequently _____ Occasionally _____ No _____
8. Do you have limited food storage space in your home? Yes _____ No _____
 Of the above food items which ones would you prefer to buy in larger sizes than you currently do if you had more room? (Give housewife attached list.)
9. How many people do you buy food for? _____

Breakdown

<u>Age</u>	<u>Number</u>
Over 18	
10-18	
0-9	

10. Do you have a budget for food? Yes _____ No _____
 If yes, for what period is your budget? Weekly _____ Monthly _____ Biweekly _____
11. How much do you spend, on the average, for food per week? \$ _____
12. Do you receive food stamps? Yes _____ No _____
 If yes, how much per week? \$ _____
13. Do you own or rent this house or apartment? Own _____ Rent _____
 What is your monthly payment on this house or apartment? \$ _____

To be filled in after leaving home

1. Estimate age of housewife _____
 2. Condition of house or apartment _____
 3. Specific comments housewife made in her purchasing habits which supplement the questionnaire _____
-

References

- Alcaly, R. & Klevorick, A. Food Prices in Relation to Income Levels in New York City, Journal of Business, 1971, 44, 40-46.
- Becker, Gary S. A Theory of the Allocation of Time, Economic Journal, 1965, 75, 493-517.
- Bureau of Labor Statistics. Prices Charged in Stores in Low and High Income Areas of Six Large Cities, February, 1966. Special Studies in Food Marketing, Technical Study No. 10, National Commission on Food Marketing, Washington, D. C., June, 1966, 121-144.
- Federal Trade Commission, Economic Report on Food Chain Selling Practices in the District of Columbia and San Francisco Washington, D. C., 1969.
- Frank, R., Douglas, S., & Rolli, R. Household Correlates of Package-Size Proneness for Grocery Products. Journal of Marketing Research, 1967, 4, 381-384.
- Goodman, Charles S. Do the Poor Pay More? Journal of Marketing, 1968, 32, 18-24.
- Hicks, J. R. Value and Capital. Second Edition, London, Oxford University Press, 1957.
- Kunreuther, H. The Effect of Quantity Discounts on Consumer Demand: An Application of Inventory Theory. Center for Mathematical Studies in Business and Economics Report No. 7009. University of Chicago, Graduate School of Business, 1970.
- Sexton, D. E., Jr. Do Blacks Pay More? A Comparison of Prices Paid For Grocery Store Commodities by Black and White Families. Unpublished doctoral dissertation, University of Chicago, 1970.
- Sexton, D. E., Jr. Comparing the Cost of Food to Blacks and to Whites--A Survey. Journal of Marketing, 1971, 35, 40-46.
- Shraier, S. A Theory of Household Behavior in Purchasing Frequently Bought Goods. Mimeograph, January, 1972.

SOCIAL MARKETING AND CONSUMERS' PREFERENCES FOR SOCIAL CONSUMPTION

James R. Bettman and Robert B. Andrews
University of California, Los Angeles

Social consumption refers to the process of goal choice and resource use at the national level. That is, as a nation we collectively set priorities for use of our scarce resources in pursuing various areas of concern such as national defense, education, health care, pollution control, and so on. This choice process has an analogue in product choices for personal or family consumption at the individual level.

There has been much study of the marketing function as it impinges upon and influences these individual product choices. Recently there also has been rising interest in the processes of social marketing and influence, through which attempts are made to influence desired social consumption patterns of citizens. Kotler and Zaltman, in a thoughtful study of social marketing, define the concept as follows (Kotler and Zaltman, 1971, p. 5):

Social marketing is the design, implementation, and control of programs calculated to influence the acceptability of social ideas and involving considerations of product planning, pricing, communication, distribution, and marketing research.

In carrying out social marketing to facilitate planned social change, two broad issues arise. First, the marketing concept, as applied to social consumption, implies that the needs and wants of citizens for societal resource use should be ascertained, to "restore consumer sovereignty in the determination of the society's product mix and the use of national resources." (Kotler and Zaltman, 1971, p. 5). The means for implementing this concept, however, are not clear. Second, questions concerning the ethics of social marketing arise. How should social marketing be carried out so that it is not inappropriately manipulative? How should the process of planning for social change, which must precede social marketing, be designed?

The present study considers the type of information and kinds of processes that would be needed to attack these issues. Little research has been done on determining the relative importance that individual citizens hold for various national or social goals. This study shows how such social consumption preference information can be gathered and presents empirical findings for national and health care goals. These goal weighings are analyzed to determine the effects of age, race, and education on such preferences. After describing the empirical results, the potential uses of this kind of data are discussed. The discussion centers on the two issues discussed above: (1) How data on preferences for social consumption can be used to improve the effectiveness of social marketing by implementing the marketing concept and by defining specific target markets; and (2) How 'ethical' social planning and marketing processes can be designed and what role data on preferences for social consumption can play. Here many ideas proposed for planning and legislation in the area of environmental protection are useful.

Measuring Individuals' Preferences for Social Consumption¹

To measure preferences for social consumption, particular societal goal areas must be designated. For the present study, data were collected for two sets of five goal areas. The first set dealt with national goals: 1) Reduce air pollution (POLL)²; 2) Slow Down Inflation (INFL); 3) Improve Racial and Ethnic Relations (RACE); 4) Decrease Crime (CRME); 5) Improve Health Care Services (HLTH). The second set was goals within health care: 1) Build More Medical School Facilities to Train More Doctors (FACL); 2) Increase Medical Research (RESR); 3) Require All Voluntary Private Health Insurance Plans to Pay for All Health Services (PRVI); 4) Make It Easy to Get Medical Treatment at Night and on Weekends (NTWK); 5) Provide Government Health Insurance for All Citizens to Pay for All Health Services (GOVI). Data on age, race, and years of education of each respondent were also obtained.

The data were collected by the Survey Research Center at UCLA as part of a large survey and comprised completed interviews from 964 households in the Los Angeles metropolitan area. The sample was a probability sample with marginal constraints on median income, percent black, and geographic area. The data were collected from November 1970-January 1971. Because of refusals and other factors for particular stimulus pairs, 787 complete sets of scale values for the national goals were obtained, and 855 sets for the medical goals.

To obtain the preference measures, data were collected using paired comparisons with an intensity rating. For the five goal stimuli in each set, ten ratings were collected from each person; one for each distinct pair of goals. For an ordered pair of goals (i, j), the individual chose which of the pair he believed was a more important goal, and then rated how much more important that goal was to him or her on a ten point scale, where zero was indifference and nine was much more important. Denote this rating for the pair of goals (i, j) by d_{ij} . We use the sign conventions $d_{ij} > 0$ if i is more important than j , $d_{ij} = 0$ if indifference, and $d_{ij} < 0$ if j is more important than i . Finally, since data on only the distinct pairs of goals are collected, we set $d_{ji} = -d_{ij}$.

The preference value for each goal held by an individual is obtained simply. For goal i , the scale or preference value p_i is given by $p_i = \sum_{j \neq i} d_{ij} / 5$. Note that by their definition, the p_i sum to zero. Thus we have a measure of the relative preference of an individual for each goal.

The data on preferences for the social goals in each set were then analyzed using multivariate analysis of variance (MANOVA)³ with three factors: age (5 levels), race (3 levels), and education (3 levels). The levels for each factor are shown in Table 1. In MANOVA, the dependent variable is a vector of values rather than a single number, as in ANOVA. In this case, the values for each goal for each person are the entries in the dependent vector. However, to perform this analysis, one goal preference must be deleted from each set, since the preference scale values sum to zero. Among national goals Slow Down Inflation was deleted and among health goals Private Health Insurance was deleted. These stimuli were chosen for two reasons: relatively small variation among the factor subgroups, and a priori empirical interest of the investigators. Discriminant analyses were performed using the sums of squares for each hypothesis as an aid to interpretation.

Table 1
Group Preference Averages for National Goals

GROUP (N)		POLL	INFL	RACE	CRME	HLTH
Age	19-25 (117)	.80	-.28	.49	.45	-1.45
	26-35 (173)	.60	-.18	-.67	.47	-.23
	36-50 (258)	.21	-.19	-.82	1.32	-.52
	51-65 (134)	-.44	-.13	-1.24	1.94	-.13
	>65 (105)	.19	-.74	-2.06	2.38	.23
Race	Black (90)	-2.34	-.11	.68	1.24	.54
	Mex-Am (237)	.19	-1.42	-1.19	1.74	.68
	White (460)	.82	.30	-.94	1.00	-1.19
Educ. years	<6 (117)	.20	-1.77	-1.87	2.43	1.00
	7-12 (433)	-.10	-.01	-.91	1.27	-.25
	>13 (237)	.98	.03	-.18	.63	-1.46
Total	(787)	.27	-.26	-.83	1.25	-.43

Table 2
Group Preference Averages for Health Care Goals

GROUP (N)		FACL	RESR	PRVI	NTWK	GOVI
Age	19-25 (134)	.28	.44	-1.38	1.00	-.34
	26-35 (189)	.19	.70	-1.48	1.03	-.44
	36-50 (263)	-.03	.64	-1.41	1.45	-.65
	51-65 (150)	.00	-.37	-.99	1.46	-.10
	>65 (119)	-.38	.10	-2.16	1.68	.76
Race	Black (89)	-.49	.29	-1.82	2.31	-.29
	Mex-Am (236)	-.82	-.82	-2.08	2.62	1.10
	White (530)	.48	.91	-1.09	.57	-.87
Educ. years	<6 (117)	-1.40	-1.64	-2.57	3.33	2.28
	7-12 (470)	-.19	.40	-1.31	1.54	-.44
	>13 (268)	1.02	1.19	-1.11	.05	-1.05
Total	(855)	.02	.37	-1.45	1.32	-.26

The mean ratings for the various groups of respondents are shown in Tables 1 and 2. The MANOVA for the preferences for the national goals showed no significant interactions, and the following significant main effects: linear trends in age ($p < .0001$); and education ($p < .0001$); a quadratic trend in education ($p = .052$) and a main effect for race ($p < .0001$). For age, the results of the discriminant analysis (not presented here) show that the most important components of the linear trend effect were increasing concern with crime and with health, and decreasing choice weighting for racial relations as age increases.

For race, two independent discriminant functions were statistically significant, thus showing a two-dimensional discrimination among the racial groups. For the first dimension, health care is the most important variable for discriminating among the races, and pollution is most important for the second dimension. By examining Table 1 the decreased concern with pollution among blacks and the decreased concern among whites for health care can be seen.

Finally, for education the most important components for the linear trend were shown by discriminant analysis to be decreasing concerns for crime and health care and increasing concern for racial relations with increasing education. The quadratic trend in education was influenced largely by the U-shaped relationships for pollution as education increases.

The MANOVA for health care goals showed the following significant effects: race-education interaction ($p = .028$); linear trend in age ($p = .043$); linear trend in race ($p < .0001$); linear trend in education ($p < .0001$); and quadratic trend in education ($p = .032$). The major factors in the interaction, according to the discriminant analysis, involved the goals of research and of government insurance. For research, the averages (not presented here) show both whites and blacks have increasing concern as education increases, but the preferences exhibited by Mexican-Americans first increase and then decrease. For government insurance, both whites and Mexican-Americans showed decreasing concern as education increases, whereas blacks first increase and then decrease.

Since there was a significant interaction, main effects must be interpreted with caution. We examined only race and the linear trend in education. For race, only one discriminant function was significant, showing the goals of research and of night and weekend care to be most important for discrimination among the racial groups. For the linear trend in education, the discriminant analysis showed the decrease in concern for access on nights and weekends as education increases to be most important, but the other stimuli also contributed moderately.

The findings themselves are of interest as empirical results. For example, crime was the greatest concern of each racial group. Following crime, however, the differing concerns of racial groups for national goals were shown. Blacks and Mexican-Americans were concerned with health care, whereas whites were not, perhaps reflecting the barriers to care presented by lack of affluence and knowledge of the system. Surprisingly, blacks were the only group concerned with racial relations. For health care goals, one result was the increase of concern for facilities and research and the decrease of concern for access on nights and weekends and government insurance with education. As education level increased, individuals placed higher priority on system problems and less on their personal problems within the

system. The overall interest in medical research is surprising. Note the concern of the elderly, Mexican-Americans and poorly educated with government health insurance.

The analysis of these preferences for social consumption is interesting in its own right. However, the major thrust of this paper is the inputs such measurements can make to the social marketing process, and in particular to the two major issues discussed in the introduction. Now that it has been shown that such measures are feasible and relatively easy to obtain, the uses of such data for social marketing will be discussed.

The Use of Citizens' Social Consumption Preferences in Social Marketing and Planning

The Marketing Concept in Social Marketing

To apply the marketing concept to social goals, preference ratings of citizens for alternative goals are necessary. As in private markets, the design of the product or service requires knowledge of consumer desires and goals. The measures described above attempt to ascertain such needs and desires for social goals. These means can serve to direct the social marketer to use societal resources toward the achievement of those goals most valued by the citizen consumers. For example, if any social marketing effort for health care programs were to be mounted, the health care priority measures given above could be quite useful. In particular, the somewhat surprising results on the relative ranking of national health insurance and of allocation of funds to research would be of great importance if consumer sovereignty is to be maintained.

There is a further use of social consumption preference data that aids making social marketing more effective that is related to the idea of implementing the marketing concept. If the social marketer determines what citizens' desires are, and if he designs a 'social product' or program that is consistent with these desires, it will be easier to market that program. As Kotler and Zaltman (1971) point out, this is the essence of Lazarsfeld and Merton's (1949) idea of canalization. There is an existing base of preferences and attitudes upon which the social marketer can build. Since social marketing often deals with 'products' that affect areas where deep-rooted attitudes are found (e.g., environmental issues), it is even more important for the social marketer to have some knowledge of his target market's preference structure. Wiebe's (1952) idea of force, involving the person's predisposition toward a goal, also implies that knowledge of preferences for social consumption can lead to more effective social marketing.

Finally, analyses of the differences in the preference patterns of different subgroupings within the population of interest, such as those performed above for age, race, and education, can help the social marketer gain insights into the value structures of different target market segments. For example, the data show that blacks and Mexican-Americans value certain health care goals differently (e.g., national health insurance and research). Thus, to be effective, social marketing programs dealing with minority health care problems would have to take these differences into account before implementing any programs. However, note that the definition of target segments is not this simple. As the race-education interaction for health care goals shows, the simplistic one-variable characterizations of interest

groups often used by social planners is not sufficient. Groups must be defined by more complex patterns of variables rather than by one variable alone (e.g., by race, income, age, and education rather than race alone).⁴

An Ethic of Social Marketing

Data on preferences for societal consumption can also be useful in designing an 'ethical' social marketing process. The problem that arises in social marketing is the possibility that citizens may be 'sold' social programs that they do not really want. Resources may not be allocated according to their desires. Kotler and Zaltman express this as a possibility of manipulation (Kotler and Zaltman, 1971, p. 11). How can a social planning and marketing process be designed that avoids this problem? An attempt will be made to sketch elements that such a process should contain, and then these elements will be integrated into a process design. Finally, current progress toward this type of design will be assessed.

Elements necessary for an 'ethical' social planning and marketing process include: 1) open planning; 2) impact statements and evaluation of alternatives; 3) citizens' preferences for social consumption; 4) an agency independent of the change agency or social marketer that serves as a system entry point and information system. These elements will each be discussed in more detail.

The concept of open planning has been discussed by Borelli *et. al.* (1971) in the context of power plant siting decisions. Open planning entails a) involvement of citizen groups before major decisions are made, not after the point of no return for a project has been passed; b) open discussion and even solicitation of views, including extreme views; c) full disclosure of relevant information; d) compromise. Of course, in many social marketing arenas, the heated controversy developed in power plant siting decisions will not arise. However, the principles involved are important: (1) for social consumption patterns, public views must be heard, and (2) the social marketer must be open about his plans. Social marketing of family planning and birth control programs might be an instance where open planning would be particularly relevant.

Impact statements are assessments of the impact of a proposed program on various societal groups, the environment, health levels, etc., to the extent to which these impacts can be predicted. For any adverse impacts, alternative programs or subprograms should be described that would avoid the impact, if possible. Further, if these alternatives are not chosen, reasons for this must be given. The social marketer would be responsible for providing a detailed statement of this information for any major 'social product'. In essence, this information is an attempt to make the decision process of choice among alternatives an open decision process rather than an implicit or private one.

Preference ratings for social consumption represent a complement to, and not a substitute for, open planning. Open planning should generate some data on the views of the public. However, by its very nature, open planning is a forum of limited size, and hence gives a potentially biased view. A survey methodology such as that used in this study provides a mechanism for sampling the views of various interest groups in much more depth. Of course, such information sacrifices the implications that can possibly be derived by deeper probing in open planning. There is a tradeoff between breadth and

depth. Finally, there is always the thorny question of the extent to which the public's priorities should govern in programs. That is, it may not always be appropriate to follow 'public wisdom'.

Kotler and Zaltman (1971, p. 10) described the change agency or social marketer as having a research unit that collects information to feed into a planning unit. It is the contention of the present study that there should be instead an agency independent of the change agency. This is for two reasons: (1) to avoid conflicts of interest with public interest groups and insure full disclosure of relevant information and (2) to serve as a highly visible entry point for interest groups into the system. Consumer groups are often frustrated, particularly in social marketing situations, by overlapping jurisdictions and complex governmental structure, so that the group does not know how to enter the system (this is often particularly true of health care and environmental problems). Having such an independent agency would serve Wiebe's functions of direction and mechanism (Wiebe, 1952). The independent agency would be responsible for gathering the preference data discussed above and also technical data needed to assess impacts.

The social planning and marketing process would then function as follows: the change agency would be continually apprised of social consumption preferences by the information agency. If the need for a new 'social product' to be socially marketed was discerned, open planning for this product would be instituted. Normally this would be initiated by the change agency. However, it is possible that the information agency, in its function as an entry point, could ask the change agency to open plan on a social product desired by an interest group that had entered the system in this manner. As plans progress, impact statements and alternatives would be proposed. At each step of the process, monitoring of the larger public's social consumption preferences would be used in conjunction with open planning inputs to guide further action.

Such a process certainly takes time and resources. However, we have entered a stage of societal awareness and concern where such processes for 'social products' may be necessary. It is "decision-making in the face of paradigm changes, when new fundamental assumptions enter the picture." (Borelli et. al., 1971, p. 19). Efficiency, even efficient social marketing, can no longer be the sole criterion. Pragmatically, however, resources are limited. Every social program probably cannot go through this extensive process. Only major programs should be accorded selectively this extensive consideration. Other social marketing endeavors could be dealt with in a more abbreviated fashion by using only the social consumption preferences supplied by the information agency. Finally, it would seem that there should be a single information agency or possibly one such agency for each of a few major problem areas.

What actual progress has been made in implementing such a system design in practice? Most progress in this area has been made in the area of environmental quality decisions. Let us take each of the system elements in turn. Open planning has been attempted in a few cases. Northern States Power Company in Minnesota developed a Citizens Advisory Task Force that essentially participated in open planning for a power plant siting decision (Borelli et. al., 1971, pp. 36-39). Impact statements are required for major Federal actions affecting environmental quality by the National Environmental Policy Act of 1969. There is to be "a detailed statement by the responsible official on:

(i) the environmental impact of the proposed action,

.

.

(iii) alternatives to the proposed action. . ."

[Pub. L. No. 91-190, Title I, § 102 (C), 83 Stat. 853 (1970) codified at 42 U.S.C. §§4332 (1970)].

In addition, recent court cases⁵ have changed burden-of-proof rules to essentially demand open decision making in environmental cases. The principle espoused has been well stated by Krier (1970, p. 115):

Once a reasonable showing is made that a proposed course of action poses a probable threat of significant environmental damage, the body desiring to initiate that action should be required to come forward with evidence on the likelihood of such damage, the unavailability or unfeasibility of alternatives, and the justification for its activities.

Flack and Summers report the use of priority data for water quality goals in resolution of conflict for water resources planning (Flack and Summers, 1971). They made ingenious use of interactive man-computer systems and computer graphics to resolve conflicts among decision makers. In particular, conflicts between those desiring a social marketing product and citizen representatives could be attacked.

With regard to the idea of an information agency and entry point, Krier reports that the California Environmental Quality Study Council performed many of the same functions during its first year of operation as those proposed for the information agency (Krier, 1971). The Council actually went even further and became nearly a public ombudsman and advocate for environmental affairs. Finally, the most extensive total system design that involves many of the ideas noted above is the proposed design of a system for the government of Puerto Rico involving citizen feedback, open planning, and an information agency (Chandler, 1970).

In conclusion, gathering data on citizens' preferences for social consumption has been shown to be both feasible, as evidenced by the empirical study of this paper, and highly central to social marketing. Implementing the marketing concept and designing an 'ethical', non-manipulative social marketing process needs such inputs. Since the proposed social planning and marketing process has used examples mainly drawn from the environmental quality arena, it may seem somewhat overdrawn for less heated social marketing areas. However, the principles espoused are just as relevant for such areas. Openness in decision making and acknowledging citizen preferences are necessities in the present and foreseeable social climate of public awareness and concern.

Footnotes

1. This study is based upon data collected under U.S.P.H.S. Grant 95-505 for training in Comprehensive Health Planning.
2. The abbreviations listed after each goal are used in the data summary tables.
3. Analyses were performed using the program Multivariance, written by Professor Jeremy D. Finn, State University of New York at Buffalo.
4. This type of result has also been found in a different context of attempting to relate personality variables to product use behavior (Sparks and Tucker, 1971).
5. In particular, the 'Wildlife Preserves' case, Texas East. Trans. Corp. versus Wildlife Preserves, Inc., 48 N.J. 261, 225A. 2d 130, 137 (1966).

References

- Borelli, Peter, Mahlon Easterling, Burton H. Klein, Lester Lees, Guy Pauker, & Robert Poppe. People, Power, and Pollution. Report Number 1, Environmental Quality Laboratory, California Institute of Technology, September 1971.
- Flack, J. Ernest & David A. Summers. Computer Aided Conflict Resolution in Water Resource Planning: An Illustration. Water Resources Research, 1971, 7, 1410-1414.
- Kotler, Philip & Gerald Zaltman. Social Marketing: An Approach to Planned Social Change. Journal of Marketing, 1971, 35, 3-12.
- Krier, James E. Environmental Litigation and the Burden of Proof. Law and the Environment, 1970, 105-122.
- Krier, James E. Environmental Watchdogs: Some Lessons from a 'Study' Council. Stanford Law Review, 1971, 23, 623-675.
- Lazarsfeld, Paul F. & Robert K. Merton. Mass Communication, Popular Taste, and Organized Social Action. In Wilbur Schramm, (ed.), Mass Communications, second edition (Urbana, Illinois: University of Illinois Press, 1960), 492-512.
- Sparks, David L. & W. T. Tucker. A Multivariate Analysis of Personality and Product Use. Journal of Marketing Research, 1971, 8, 67-70.
- Stevens, Chandler H. Science, Government, and Citizen Feedback. Operations Research, 1970, 18, 577-591.
- Wiebe, G. D. Merchandising Commodities and Citizenship on Television. Public Opinion Quarterly, 1951-52, 15, 679-691.

THE INDEX OF CONSUMER SATISFACTION:

METHODOLOGY

James C. Lingo, University of Michigan
 Martin Pfaff, Wayne State University and University of Augsburg

I. Introduction

1. Subjective Market Failure and Consumer Dissatisfaction

The intellectual landscape of the past decade has been marked by dramatically contrasting signposts: On the one hand, economic measures of the growth of national product and income have pointed to a society of affluence. On the other hand, the storm signals about urban unrest, social disintegration, and widespread poverty in urban ghettos and rural enclaves has suggested that there are certain basic features of American society which cry for remedy.

In the area of economic and market performance, we notice a growing dissatisfaction of consumers with goods and services offered in the market place. This may be due more to a rising level of awareness fostered by news media than by any deterioration in the objective conditions under which these goods and services are being provided. Nonetheless, if consumer satisfaction is viewed as the objective of market actions in an open society, it has to be taken into account when formulating public and corporate policy.

The economy theory of "market failures" or "market inefficiencies" recognizes that deviations from an economic ideal of welfare maximization can be brought about by (a) monopoly forces which distort prices and output levels; (b) indivisibilities or "lumpiness" in production or consumption which make impossible the fine marginal adjustments postulated by welfare theory; (c) increasing returns to scale which obviate competition since a single firm can supply a market at a cost lower than several competitors; and (d) externalities in production or consumption which lead to a divergence between private and social benefits and costs. To these we may add market imperfections due to limited information, cyclical instabilities, and so on. These theories, however, assume implicitly that diminution of social welfare occurs because of forces largely outside the individual consumer. In fact, the consumer sector is treated largely as a more or less homogeneous group.

Such a view neglects, however, the wide variance in the degree of (dis) satisfaction among different consumers. What is needed therefore is a theory of "subjective" welfare: Dissatisfaction of the individual consumer results from the discrepancy between ideal and actual attribute combinations. Such a discrepancy, moreover, can result from both outside and within the consumer: If, for example, the perceived ideal changes, dissatisfaction can go up even if the so-called objective circumstances do not change at all! This view of subjective market failure treats the consumer group as highly heterogeneous.

The Index of Consumer Satisfaction is a measure of this subjective welfare of consumers as influenced by the attributes and circumstances under which goods and services are offered in the market. In this paper we describe the methodological foundations of (a) representing market goods and services in terms meaningful to the subjective experience of the consumer (the representation problem); (b) predicting overall satisfactions (the composition problem);

and (c) aggregating individual satisfaction scores into group satisfaction and potentially into an overall satisfaction measure for the nation (the aggregation problem).

2. Measurement of Economic and Market Performance

It is perhaps the conflict between overall progress signified in economic accounting systems, and social disharmony and misery as expressed by other measures on the social plane, which has motivated the growing concern and the re-examination of our existing measures of economic and social performance.

Institutes, conferences and professional meetings have been addressed to the question of whether the national economic accounts are deficient in any way and how they could possibly be improved to reflect some of these economic and social problems. The Conference on Research in Income and Wealth, (sponsored by the National Bureau of Economic Research at Princeton, New Jersey, November 4-6, 1971) is one of the more recent examples of this search for new ways of measuring economic and social performance. One of the major themes that came out of this conference was the need to view the performance of the economy and of society in terms of the positions of specific sub-groups and social sub-strata of total society rather than in an overall fashion. Furthermore, the need to recognize social costs in the framework of national accounts that result, for example, from the deterioration of the environment, is widely felt. No doubt, the presently used framework of national accounts has taken note of the government in the economy; however, it is largely a market-based system and it neglects the wide role of non-market transactions.⁴ Others again have pointed to the need to recognize the various imputations, such as imputed rent on owner-occupied dwellings, the services of housewives, consumer durables, etc. in the national income accounts explicitly (Ruggles & Ruggles, 1970).

Even if all of these proposals for the extension of economic measures of the economic performance of our society are adopted, they fall short of describing some significant aspects of the quality of life and of social existence. The reason is simply that such measures rely on dollar measurements of transaction flows; they reflect social problems such as discrimination, inequality, crime, etc., through social gains or losses which would be added on to the economic gains or losses reflected in the national income accounts.

This recognition, together with an increasing awareness of the social problems of our society, have increased the level of interest in a system of social indicators which would be a more direct reflection of the areas of social concern. No doubt there are conflicting views as to what social indicators should do and how they should relate to other measures of economic and social performance in our society. Most would agree, however, that social indicators should be relatively comprehensive; and that they should relate to a system of overall social accounts which offers perhaps some parallels to the presently used system of economic accounts.⁵ This point of view, however, is bound to be somewhat controversial at a time when a general theory or model of how social processes relate to economic process has not yet gained wide acceptance.⁶

Several measures have been employed or proposed to assess the performance of the market system or of its components. The type of measure employed depends on the theory or model of the market that the investigator holds either implicitly or explicitly (Pfaff, 1968).

Within any of these approaches, the process of measurement is influenced by the concept of market performance that the investigator adopts: He can evaluate the system by imposing his own "external" values; or else, he can observe the value patterns prevalent in the system itself and judge its performance by these "internal values" (Douglas, 1968; Cook, 1968). In any case, he is forced to embrace some kind of norm or yardstick, without which measurement is impossible. Even if he falls back on the traditional economic norm of "efficiency" or productivity, he is no less normative: In the face of conflicting social goals, the selection of any one norm represents a normative act par excellence.

Similar considerations apply to the measurement of the performance of the non-market sector, including the public economy.

Among the measures of market performance we find the 1. "value-added method", 2. physical productivity analysis, 3. flow analysis, 4. market structure analysis, and 5. the application of welfare economics (Buzzell, 1959; Pfaff, 1969).

3. The Measurement of Utility and Welfare

The second major approach to the problem of assessing the performance of the economy and of society tends to get at the very end or objective of the whole economic and social process which, in an open-society, is generally held to be the satisfaction of the needs of the population. No doubt, this focus has a long tradition in economic theory where the utility or the "want-satisfying power" of a commodity or service was the central focus of the theory of value. Indeed, if the utility of various economic and social acts could be assessed empirically, a system for the evaluation of performance par excellence would be at hand. Unfortunately, while the concept of utility has been the main foundation for the theory of value, it has not really been extended to practical social policy.

Ever since Robbins' An Essay on the Nature and Significance of Economic Science (Robbins, 1932), it was held that value judgments should not be part of "objective" or scientific analysis. This position is reflected in the work of Kaldor, Hicks, and others who sought to formulate a value-free New Welfare Economics (Kaldor, 1939; Hicks, 1939, Reder, 1947). However, their approach contained ethical assumptions like consumer sovereignty, as was pointed out later by Bergson (1938) and Baumol (1946-47). Bergson proceeded to explicate the role of value judgments in the form of his Economic Welfare Function. Assuming that all non-economic, i.e., social or institutional variables (like government and social structure) remain constant, economic welfare is determined simply by economic variables; or, it is assumed that changes in the values of economic variables are not large enough to bring about changes in the institutional environment. Thus, "The Economic Welfare Function simply traces the effects on the individual's welfare of changes occurring in one part of the environment, the other portions remaining unchanged" (Rothenberg, 1961, p. 10).

Social welfare functions have been applied in areas where alternative policies were to be evaluated in the face of conflicting ends. Many criticize however the basis on which these judgments are formulated. In their view the de-facto policies of government cannot necessarily be taken as an expression of a social welfare judgment.

Others again would reject the concept of welfare function on theoretical grounds, based on Arrow's celebrated Impossibility Theorem (Arrow, 1951). A reinterpretation by Rothenberg appears to offer a way out of this theoretical impasse (Rothenberg, 1961).

In the face of the problems associated with the measurement of utility, several scholars have turned to an attempt to measure welfare rather than utilities. No doubt, the concept of "measurable welfare" is somewhat different from utility, but it may provide the very alternative which is required for public policy making. One of the most interesting of these attempts has been the work of the United Nations Research Institute of Social Development which has consistently attempted to formulate indices of the level of living, of welfare, and so on (UNRISD). In one of these reports, Jan Drewnowski defines measurable welfare as "a quantitative expression for the satisfaction of needs of a population," (Drewnowski, 1968). He intends to formulate such a measure that should present a macro-concept which is fully measurable in cardinal terms, summable, and comparable interpersonally, internationally and through time. Analogous to the concept of utility, he interprets measurable welfare as the dependent variable of a preference function:

The Level of Welfare is the sum total of characteristics of the person or the population expressed in quantitative form and referring to the state of satisfaction of the needs of that person (Drewnowski, 1968).

Drewnowski proposes to measure the level of welfare indirectly by relating different component measures which are expressions of their somatic status (physical development level), to the educational status, and the social status. He touches on the problem of weighting: He points out that in the absence of information to the contrary, the various sub-components would be given equal weights within an overall component, and various components will be again equally weighted for a total "level of welfare" index measure. The kinds of indicators he would use are essentially cardinal measures of the health status expressed, for example, through life expectancy, or percent distribution of healthy versus sick people. Similarly, he would measure educational status by literacy rate, educational attainment rate, and so on. In the area of social status he attempts to find numerical expressions for activities and social trends which would have profound ramifications for the level of welfare of a given society; among these he lists the level of integration within the nation, social groups, and family, as well as the level of participation of various social actors.

4. "Subjective" Measures of Economic and Market Performance

As an alternative to these indirect and "objective" measures of welfare we proceeded to measure the level of satisfaction of the needs of individuals and persons not by some abstract indicator but on the basis of an individual's own perception of his own satisfaction. This approach then amounts to a direct and "subjective" approach to the measurement of welfare, based on the respondents' own evaluation of their satisfaction with particular aspects of goods or services provided by the market or the public economy (Pfaff & Pfaff, 1969). We attempted to measure the performance of the market economy

by what we have termed an Index of Consumer Satisfaction which measures a cross section of the population's evaluation of the various services provided by the market-distribution system. Furthermore, we measured the performance of the public economy, that is, the services provided by the public sector, by the citizens' evaluation of the attributes of such services and the relationship of these attributes to their own needs and expectations. This Index of Citizen Satisfaction has only been explored initially, and a pilot test of this Index is contemplated for the near future. These indices of consumer and citizen's satisfaction are, therefore, social indicators which attempt to measure the human satisfaction which results from the operations and performance of the market and public economy. They aim at judging the quality of life produced by various products and services available to consumers through market offerings on the one hand, and by the various programs of government which have economy and noneconomic consequences for different groups in society, on the other. The ultimate aim of these indices is to provide measures of the satisfaction with the different aspects of life experienced by consumers and citizens. A major interest in this study is the specific and general aspects of consumer dissatisfaction or citizen dissatisfaction with the quality of life, and how such satisfactions or dissatisfactions are distributed across specific racial, ethnic, age, income and other groups, defined on the basis of psychological, social, economic and cultural criteria.

Subjective measurement of satisfaction has a long tradition at the hands of industrial psychologists who were concerned with the more specific area of job satisfaction. Their aim was to measure the relationship between satisfaction and productivity. However, "no really substantial, reliable or general correlation between satisfaction and productivity has been established" (Smith, Kendall, & Hulin, 1969, p. 3; Brayfield & Crockett, 1955; Smith & Cranny, 1968).

These measures of work or job satisfaction are concerned with the input side of the economic process. At the other extreme, we have what may be construed as measures of the output of the social system, viz. the happiness with life in general, and with specific facets of life in particular. Among these, Hadley Cantril (1965) studied the happiness, and the hopes, and fears of persons in fourteen countries, Bradburn, (1969), and Caplovitz, (1965) and Wilson (1967) discuss the distribution of avowed happiness across different groups. Inkeles (1960) concluded that individuals of higher socio-economic status tend to be more satisfied with their lives than those of lower status. More recently, Easterlin (MIMEO) synthesized the findings based on cross-section analyses of individuals' self-assessment of happiness. He inferred a strong positive relationship between happiness and income within a society, but noted the absence of such a relationship across countries at a given point in time. Gurin, Veroff, and Field (1960) study happiness in the context of the overall mental health of Americans.

There has been no study, however, of the satisfaction of consumers with the performance of specific aspects of the market system on the one hand, and there are only a few approaches to the study of citizens' satisfaction with public goods provided by the government, on the other (Stagner, 1970).

The methodological foundation for the computation of these Indices of Consumer and Citizen Satisfaction was provided by nonmetric scaling techniques, as developed by Guttman, Lingoes, and others.

The computation of these indices was based on a set of procedures appropriate to the assumptions which we felt were justified by the data collected. In so doing we aimed at achieving reliable, valid, and univocal scales. We shall now turn to a description of the methods employed and their underlying rationale.

II. Research Strategy

We impose the logical requirement that individuals or groups having higher values than others on avowed overall satisfaction are also more satisfied (or, at least, see or report themselves as being more satisfied) in respect to all the components that determine the overall satisfaction index. Formally, we are imposing a homogeneity criterion on our indices, such that any pair of individuals/groups which are nearer one another than either are to any other individual/group (in terms of their indices of satisfaction), will have profiles over the components of the index which are closer together. The index is, therefore, predictive of its constituents.

To the extent that a set of variables is factorially complex (requiring more than one dimension to explain the inter-relations) we have a partial order defined on the individuals' scores. This implies that a given person may be higher than another on one dimension but the reverse may be true on another dimension. Take, for example, the case of one baseball player being a better hitter than runner and the reverse for another player; how can we compare these hypothetical players as to baseball talent? Can we arrive at some measure to make the following statement sensible: "Lou has more baseball talent than Joe."? In essence, can we integrate or map a complex domain onto a line so that the order of scores is meaningful? Our work-a-day world in which decisions must be made (e.g., salaries to be paid and honours to be bestowed in the case of our baseball players) requires that we derive some "rule" which will permit us to compare the incomparable with a minimum loss and maximum sense. Many of the techniques to be described below had their primary justification for use in their ability to fulfill this goal.

In contrast with such measures as Drewnowski (1968) used, such as life expectancy and percent distributions of healthy versus sick people, our "measures" of subjectively perceived (or reported) satisfaction with various product attributes, products, and product classes cannot be assumed to have the same statistical properties or measurement status. As an example, take the following five-point scale:

Very	Somewhat	Neither	Somewhat	Very
Satisfied	Satisfied	Satisfied nor	Dissatisfied	Dissatisfied
		Dissatisfied		

What numbers shall we assign to the above various verbal descriptions? Some would say that the integers 1 to 5 would be appropriate, while others with equal justification might wish to assign the following number sequence: 1, 3, 4, 5, and 7. Each of these assignments entails assumptions about the relationship between language and a number system. Each refers to some hypothesis regarding the regression system which relates these values among themselves and to other variables (which takes place in the context of

establishing functional or statistical relationships). The 1-5 assignment assumes equal distances between adjacent points, while the second correspondence assumes equal distances for only part of the scale (the middle) and insists that the end points are twice as far from their immediate neighbors as are points in the middle of the range. Both, in this instance, assume a symmetric scale. Other assignments are possible and need not be restricted to linear scales, e.g., 1, 2, 5, 2.5, and .5, for the case of a nonsymmetric and curvilinear scale. In any event, every such assignment carries implications regarding regression estimates (Guttman, 1971) whether we are aware of them or not. If we do not presuppose specified properties of the scale (e.g., the exact spacing between points), then various techniques become applicable for deriving this spacing. We are then involved in the process of scaling, which is the subject matter to be discussed under the topic of scoring methods.

1. Dimensional Simplification

A number of techniques are available for studying a set of interrelationships and their patterning. These may be broadly classified under the headings of metric and nonmetric methods, and each can be further subdivided into dimensional and adimensional or typological (clustering) procedures. The chief difference between the class of metric and the class of nonmetric methods lies in the level or quality of the data collected, or, better yet, it rests on the level which we assume to exist, since the numbers themselves are "dumb" when it comes to communicating their origin or use. If we use metric procedures we assume that the difference between, say, a correlation of .4 and .5 is equal to that between .8 and .9, and we judge the goodness of fit for the solution obtained in respect to the degree to which such equalities are preserved. A nonmetric approach, on the other hand, is based on the assumption that only the order of these relationships is meaningful. It would consider any solution which would be order-preserving as being satisfactory. For example, .2, .25, 1.3, and 2, corresponding to the above four correlations seriatim.

The preceding distinction can be characterized in terms of the concept of a loss function, which is a measure of the errors we are willing to tolerate for whatever purposes we have in mind in analyzing the data in the first instance. Implicit in many metric techniques is the minimization of the loss function:

$$(1) \sum_{j>i} (\hat{r}_{ij} - r_{ij})^2,$$

whereas, the comparable nonmetric loss function would be:

$$(2) \sum_{j>i} (\hat{r}_{ij} - r_{ij}^*)^2,$$

where R^* is a monotone transformation of R (the input correlations), e.g., \hat{R} permuted into the order of R , where \hat{R} is the matrix of reproduced coefficients based on m dimensions.

It might be obvious to many that loss functions play a critical role in constructing measures of central tendency in statistics, or that they are implicit in everyday situations where some loss might be incurred as a result of making the wrong guess (e.g., in gambling). But most are unaware of their implications in the field of multivariate analysis. This may be true partly because the errors involved do not have clearcut costs associated with them. A gambler knows that if he guesses more often wrongly than correctly, he will be out-of-pocket. The researcher, by insisting on an inappropriate loss functions, generally pays in the coin of complexity. Our preference for nonmetric over metric procedures is based on the fact that the former greatly simplifies our interpretation of patterns. It results in fewer dimensions, in general, and the assumptions for its use are more consonant with the types of data collected by social scientists.

Both dimensional and adimensional approaches are useful under given conditions. If the pattern of coefficients is such as to result in disjoint subsets, then the adimensional or typological approach is appropriate, otherwise, the dimensional approach is in order. Without analyzing the data, however, one cannot usually say in advance whether one's approach should be dimensional or adimensional. Clustering procedures used in tandem with dimensional analyses are often most productive, as long as some decision rule is present for rejecting one or the other hypotheses. The typical clustering procedure, like item analysis, always permits one to admit the hypothesis. There are clusters or items for constructing a scale. This state of affairs is, of course, undesirable, and we offer one possible solution to correct it. Dimensional analysis presupposes a continuous space in which points or vectors may lie and, as such, can be given a geometric representation. Adimensional analysis, on the other hand, is based on the assumption of discontinuity; it has its natural representation in terms of unordered classes. A third possibility, lying between these extremes, are metric spaces, which in fact (but not in principle), have dense regions of points scattered throughout it (like the heavens) and, as such, may be fruitfully viewed as consisting of ordered clusters.

The following three nonmetric techniques have been found useful in studying data structures derived from satisfaction ratings, from the point of view of both interpretation and as a desirable preliminary to constructing homogeneous indices: a) monotone distance analysis (Guttman, 1968; Roskam & Lingoes, 1970; Lingoes & Roskam, 1971); b) monotone vector analysis (Lingoes, 1966; Lingoes & Guttman, 1967); and, c) probability evaluated partition analysis - a clustering procedure (Lingoes & Cooper, 1971). We shall briefly discuss each of these methods in turn.

a) Monotone Distance Analysis

Let us start with a given arbitrary symmetric matrix of coefficients, all of which are assumed to be comparable; this would be the case, for example, with correlation coefficients. The goal of monotone distance analysis is to determine the smallest space in which the points lie, such that the distances among the points are isotonic (order isomorphic) with the order existing among the coefficients. Short of a perfect order-preserving, monotonic transformation of the input values, we seek the best possible fit for a stated number of dimensions, by minimizing the loss function:

$$(3) \sum_{j>i} (d_{ij} - d_{ij}^*)^2,$$

where D is the generalized Euclidean distance in m dimensions and d^* is a monotone transformation of those distances in terms of the order present in the input matrix R , for:

$$(4) d_{ij} = \left| \sum_{a=1}^m (x_{ia} - x_{ja})^2 \right|^{1/2},$$

where X is a n (points) \times m (dimensions) rectangular coordinate set of real numbers. If we are dealing with similarity data, such as correlations, then small distances correspond to large correlations, e.g., $d_{ij} < d_{kl}$ whenever $r_{ij} > r_{kl}$ (for all i, j, k , and l ; $i \neq j$ and $k \neq l$). On the other hand, if our coefficients measure dissimilarity, then $d_{ij} < d_{kl}$ whenever $r_{ij} < r_{kl}$. Our geometric representation reflects the order information in R and we interpret the space in terms of proximity. Points which lie close together in the space have something in common. On the basis of what we know about the variables and our hypotheses, we attempt to determine the law of formation giving rise to the particular observed configuration (Green and Carmone, 1970). A short example will illustrate some of the principles.

Suppose we had the following correlation matrix for five variables:

```

X
.02 X
.00 .02 X
-.02 .00 .02 X
-.04-.02 .00 .02 X

```

Can we find a nonmetric solution in terms of a set of distances which will strictly preserve the above order? According to the Guttman-Lingoes theorem (Lingoes, 1971), $n-2$ dimensions will always satisfy our nonmetric requirement and such solutions are neither informative nor unique. Our question should rather be, can we find an order-preserving solution in fewer than $n-2$ dimensions? The answer, of course, is "yes", since the matrix was constructed to yield a one-dimensional solution. This implies that the coordinates: (0,1,2,3,4) result in the distance matrix:

```

X
1 X
2 1 X
3 2 1 X
4 3 2 1 X

```

which is inversely monotonic with the above correlation matrix.

Serving as a contrast is the metric solution for the above correlations, i.e., (-.2, -.1, .0, .1, .2), for dimension I and (.0, .1, .2, .1, .0), for dimension II, reproducing the exact values in R. In general, a non-metric model will require fewer (usually much fewer) dimensions to reflect the pattern in R than will a metric model; thus, the process of interpretation is facilitated. It will be noted in this artificial example that the non-metric solution also preserves equality for intervals. But such is not the general rule, since any monotone transformation (and not merely linear as in the present example) is equally good from the nonmetric point of view. For less than perfect fits, however, the particular monotone transformation which reduces dimensionality tends to be unique. As the number of points increase relative to the number of dimensions in which a solution is sought, the nonmetric order constraints act as if they were metric in nature.

b) Monotone Vector Analysis

In contrast to the distance model described, we can define a set of scalar products:

$$(5) \hat{r}_{ij} = \sum_{a=1}^m x_{ia} x_{ja},$$

which will minimize loss function (2) above. Rather than determining distances which are monotone with R, we are finding vectors in the space so positioned that the product of their lengths and the cosine of the angle subtending them will yield a set of values monotonic with the input values (which are generally correlations or covariances). If we had employed (5) with (1), of course, we would be seeking a metric solution, i.e., we would be doing a factor analysis. Since most researchers have some familiarity with factoring procedures, monotone vector analysis serves as a relatively painless introduction to the area of nonmetric techniques.

Didactic motivations apart, however, there are some matrices for which a vector solution is more parsimonious than one based on the distance model. Spearman's unit rank hierarchy, for example, yields one dimension when decomposed by either the metric or nonmetric vector models, whereas n-2 dimensions are required with the distance model. There are other matrices for which the opposite holds and some yield equally parsimonious solutions for the vector and distance models. In no instance, however, would a metric solution be more parsimonious than an appropriately applied nonmetric procedure. The differences in dimensionality obtained by different models is in a large part due to communality issues. This is a topic which we can but mention rather than discourse upon here (Guttman, 1967). We can, however, put forth some guidelines for selection of models.

In general, if one has some particular composition model in mind then the method of analysis should be in agreement with that model. For example, similarity/dissimilarity ratings are consonant with the concept of closeness or proximity. Therefore, a monotone distance analysis is appropriate to such indices. Correlational data, on the other hand, quite naturally evoke the concept of covariation; as a consequence, a monotone vector analysis is the method of choice (Coombs, 1964; Roskam, 1968). In the absence of a composition model or theory about the data collected, one is free to select that method of analysis which most parsimoniously and clearly reflects those features of the data (e.g., patterning) which are to be

interpreted. The foregoing would appear to some as a cavalierly stated dictum; it might require one to perform several analyses before attaining a satisfactory representation. If the method leads to some new insights, this may be justification enough for one's pragmatism.

There are some more specific considerations, however, in model selection. Whenever there are more than three dimensions necessary to obtain a reasonable fit when using the distance model, the vector model may offer more purchase or interpretative leverage on the solution; it also provides a means for arriving at scores for individuals. Other factors may affect dimensionality. For example, choice of coefficient or non-monotone transformations of the coefficient, such as using r^2 rather than r , or even reflecting negative r 's to make them positive when applying the distance model, fall among these. But we shall not dwell on these issues here. Suffice it to say that it is important not only to match method and mode; it is equally important to understand, prior to their analysis, what the coefficients mean either in terms of proximity or covariation. If one does not take such considerations into account, one's comprehension of the configuration or geometric representation will be limited (Lingoes, 1970).

As an aid in comprehending spaces of large dimensionality or as a method for analyzing data assumed to be basically discontinuous or "clumpy", we shall now briefly outline another nonmetric technique.

c) Probability Evaluated Partitions

Let us start with a matrix of either proximity or concomitance (covarying) data, all of whose values are deemed to be comparable. The aim of probability evaluated partition analysis (or PEP) is to determine the maximum number of disjoint sets which are consistent with the hypothesis that their emergence or existence would not be reasonably expected if the space in which they could be embedded were indeed continuous. That is to say, we are proposing the null hypothesis that the space is continuous or that the points are non-clusterable, against the alternative hypothesis that the space is discontinuous or that the points are clusterable. Implicit in this formulation is the concept of statistical inference, where it is assumed that the probability is 1/2 that any given coefficient is above or below some specified cut-off value. The essential flavor of the procedure can be rather quickly captured by providing an example for analysis. Suppose we had the following matrix of coefficients and wished to test the hypothesis of continuity:

		<u>PROXIMITY MATRIX</u>								
		<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	<u>8</u>	<u>9</u>
1	X									
2	20	X								
3	15	16	X							
4	19	0	29	X						
5	2	4	30	26	X					
6	13	14	1	2	3	X				
7	17	18	4	5	6	20	X			
8	-1	6	7	8	9	22	21	X		
9	3	1	10	11	12	30	32	35	X	

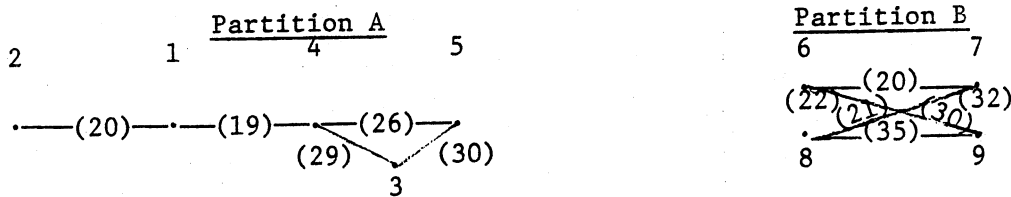
Our first step in decomposing the above matrix is to assign rank-order values to the above coefficients (from 1 to $\binom{n}{2}$, where 1 is assigned to the largest value if we are dealing with dissimilarity data or to the smallest value for similarities). Considering the relationships among these nine points to be similarities results in the following matrix of rank-order values:

RANK ORDER MATRIX

	1	2	3	4	5	6	7	8	9
1	X								
2	27.5	X							
3	22	23	X						
4	26	2	32	X					
5	5.5	9.5	33.5	31	X				
6	20	21	3.5	5.5	7.5	X			
7	24	25	9.5	11	12.5	27.5	X		
8	1	12.5	14	15	16	30	29	X	
9	7.5	3.5	17	18	19	33.5	35	36	X

, where fractional values denote tied ranks.

Our next step involves assigning zeroes to our rank matrix in terms of the above rank order until that point is reached where the matrix becomes disjoint. This implies that there is a partitioning into two sets, such that all members in one set are related to at least one other member in that set and no members from one set are related to members of the other set. In the above example, we find that 25 zeroes must be assigned before we can get the desired separation, resulting in the partition consisting of points (1,2,3,4,5) and another having members (6,7,8,9), which can be displayed graphically in terms of the original coefficients as:



It will be noted that the largest coefficient between any member of set A and set B is 18 (corresponding to the rank-order value of 25), the element in row 7 and column 2 of the above matrices, respectively. All links in A and B are greater than 18, (those points having links ≤ 18 are not graphed). Partition B can be differentiated from A in terms of the number of connections or links relative to the maximum number of connections for an order n graph, i.e., the coefficient of tightness for A is $5/10$ or $.5$, while for B this coefficient is $6/6$ or 1.0 . No attempt was made in the graphs shown above to make distance proportional to the input coefficients, although this might be possible in this instance.

The only novel feature in the algorithm to this point is the use of the ordinal rather than metric information for creating the division. The final, but crucial, step involves attaching some probability value to the partition, which would allow us to either terminate or continue the process of partitioning. Based on the Cooper-Lingoes decomposition theorem for computing probabilities of labeled disconnected graphs, we find $p = .322$ for 25 or fewer deletions (zeroes) resulting in a disconnected graph. If we had chosen an alpha level of .05 for testing the null hypothesis, we would have to accept the hypothesis of continuity, i.e., this matrix is non-clusterable. If p were $< \alpha$, we would have a basis for subpartitioning each of the partitions obtained until no further partitions resulted having acceptable p 's. The clusters obtained from PEP are invariant up to a monotonic transformation of the input values.

The preceding three methods were used for understanding the observed relationships among satisfaction scores and as a preliminary for deciding which variables to use in constructing our scales. We employed them since the scaling procedures are sensitive to lack of homogeneity among the items scaled and the process of aggregation over variables introduces ambiguity of interpretation whenever heterogeneity exists.⁷

We shall now turn our attention to constructing scales and indices based upon relatively homogeneous subsets of items.

2. Scaling

We shall be speaking of two different ways of arriving at scores for individuals: a) raw scores (RS) on satisfaction ratings, which assumes that the 7 positions on the scale are equally spaced and are orderable in terms of most to least; and, b) a second system for obtaining scores which are optimal, subject to the explicit constraint that they also be monotonic with RS, was used which we shall call optimal monotonic scores (OMS). Such a scaling stretches the intervals between scores, but maintains their order, so that either the average correlation among the set of variables analyzed will be maximal under the monotonicity restriction (Lingoes, 1972) or a given variable is best predicted by a set of independent variables (Lingoes, 1972, In Press). By scaling the data in this manner we are in effect saying that we will not assume, out of hand, that the difference between a 1 and 2 rating (one unit) is the same as the difference between a 5 and 6 in rated satisfaction.

Rescaling the RS into OMS results in a metric which will be optimal from the point of view that no other scoring system will yield larger product moment coefficients of correlation among the variables so scaled or will yield a higher multiple correlation (subject to the monotonicity constraint). The main implication of this scaling is that if we wish to aggregate across variables or individuals, then we should attempt to maximize homogeneity or predictability so that the resulting scores are as unambiguous as possible.

The degree of homogeneity/predictability achieved in the OMS scales is a function of sample, sample size, the number of categories, the number of individuals falling in each of the categories (the frequency distribution), and the set of variables included in any particular OMS scaling. Some of these parameters can be controlled and have an effect on homogeneity or predictability.

These two scoring systems (RS & OMS) were studied in terms of both correlations and predictions to see to what extent which of them might be better for achieving a reliable and valid Index of Consumer Satisfaction. We shall discuss these analyses in the sections to follow in the context of the aggregation problem. First, we shall address the procedures followed in constructing optimal monotonic scores.

a) Optimal Monotonic Scores (Maximizing Homogeneity)

The present method (CM-III) for obtaining OMS is best used when the items to be scaled are dimensionally simplified, i.e., when they approach unidimensionality. This again underlines the importance of our preliminary study of data structures.

The essence of the OMS CM-III Procedure is to perform differential stretchings of the RS scale intervals, such that the product moment correlation coefficient will adequately reflect the interdependencies in one's data. The algorithm consists of two basic steps: 1) determining the least-squares predictors over a set of variables for all individuals, and, 2) permuting these predictors into the order required for each of the n variables, i.e., performing a monotone transformation. Thus, we secure a single scoring system, differing only in the order in which these scores appear for each variable (a necessary condition for maximizing a set of intercorrelations). When our loss function is minimized we will have determined the best monotone transformation of our scores. Our new scales will be invariant up to a linear transformation. We obtain equal interval scales for each of the items in the set. If there are mild nonlinearities in the n variable regression system, CM-III will "iron them out", as it were. The method will tend to yield linear regressions for every pair of variables (a necessary and sufficient condition for maximizing intercorrelations). CM-III is a direct solution to the relational simplex problem posed by Cattell (1962). If one can attain a perfect fit for one's data, then all variables will be perfectly correlated. Short of a perfect fit, however, the average intercorrelation will be maximized.

The following simple two variable problem carried through only one iteration will help the reader grasp the essentials of the CM-III OMS procedure. Consider the ten paired RS for variables x and y :

Subject	1	2	3	4	5	6	7	8	9	10
x	1	2	2	3	4	5	5	6	7	7
y	1	2	3	5	6	4	6	5	6	7
Subject										
Mean	1.	2.	2.5	4.	5.	4.5	5.5	5.5	6.5	7.

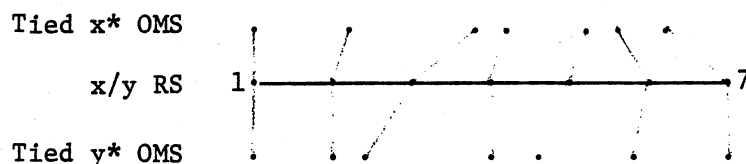
If we wished to best predict a given individual's score over a set of variables, the mean will be the best predictor (it minimizes the sum of the squared deviations). To obtain the monotone transformations of x and y , i.e., x^* and y^* , respectively, we merely permute the mean vector into the order of x and y , thus:

Subject	1	2	3	4	5	6	7	8	9	10
x^*	1.	2.	2.5	4.	5.	4.5	5.5	5.5	6.5	7.
y^*	1.	2.	2.5	4.5	5.5	4.	5.5	5.	6.5	7.

and, then, to ensure strict monotonicity, we tie the above transformed values, wherever ties exist in x and y, e.g., Ss 2 & 3 have tied x values, so:

Subject	1	2	3	4	5	6	7	8	9	10
Tied x*	1.	2.25	2.25	4.	4.5	5.25	5.25	5.5	6.75	6.75
Tied y*	1.	2.	2.5	4.75	5.83	4.	5.83	4.75	5.83	7.

As a result of just one iteration we have increased the correlation between these two variables from .845 to .915, representing about a 12.5% increase in predictability (x from y or y from x). Further iterations would, of course, improve upon this increment. The following graph shows how the above procedure performed differential stretchings on the x and y intervals:



It should be apparent from the above graph that variable x underwent more distortion than did y in achieving the increment in linear correlation.

The above technique for homogenizing regression systems can be carried out for two or more variables, which are essentially unidimensional or which cover a unified domain of behavior or attitude. We shall now demonstrate the CM-III scaling procedure on some real data, i.e., on satisfaction ratings obtained from 342 women in an initial survey we carried out under sponsorship of the U. S. Department of Agriculture. Our variables in this example are five satisfaction scales for preselected attributes, i.e., packaging, taste, nutritional value, availability, and price for the product of luncheon meats.

In table 1 below we give the CM-III OMS values corresponding to the original RS assignments 1 to 7 (after we have added in the respective means for the five attributes based on RS, since the CM-III Procedure yields values whose mean is zero and whose variance is unity). We also list the delta (increment) for the six intervals, so that the differential stretchings will be apparent.

Table 1

CM-III OMS Values for Luncheon Meats Attributes

RS	Nutritional									
	Packaging		Taste		Value		Availability		Price	
1	1.45	.93	1.48	.96	1.59	.87	1.27	1.09	2.65	.48
2	2.38	.68	2.44	.67	2.46	.69	2.36	.67	3.13	.58
3	3.06	.56	3.11	.47	3.15	.56	3.03	.48	3.71	.49
4	3.62	.46	3.58	.37	3.71	.42	3.51	.46	4.20	.48
5	4.08	.26	3.95	.36	4.13	.43	3.97	.25	4.68	.37
6	4.34	.45	4.31	.50	4.56	.49	4.22	.38	5.05	.75
7	4.79		4.81		5.05		4.60		5.80	
Mean	2.63		2.67		2.97		2.32		4.39	

Apart from the substantive findings (e.g., these women were most dissatisfied with the price of luncheon meats and most satisfied with their availability), we can see that each attribute has been rescaled by CM-III such that the intervals appear quite different from those present in the RS 1-7 scale. As a result of the above monotonic transformation we increased the average correlation among these five attributes from .427 to .442 (an increment of 1.2% in variance), which suggests, at least in this instance, that for all practical purposes we had a linear regression system and could have just as well have used the raw scores, themselves. In other instances, of course, we might achieve greater gains in homogeneity if nonlinearities were present.

In Table 2 below we display the product moment correlations for both before (RS) and after (OMS) scaling to afford the reader some insight into which variable/s was/were affected most by altering scale intervals.

Table 2

Correlations among RS (above) and OMS (below)

Item	1	2	3	4	5	Mean
1	1.00	.40	.38	.39	.38	.388
2	.43	1.00	.75	.36	.46	.492
3	.41	.74	1.00	.35	.47	.488
4	.41	.43	.39	1.00	.33	.358
5	.37	.46	.46	.34	1.00	.410
Mean	.405	.515	.500	.392	.408	

The average correlation coefficient was increased most for item 4 (Availability) and least for item 5 (Price), which suggests a mild departure in linearity of regression for the attribute of availability. The increase in homogeneity would not be deemed sufficient, however, to warrant the application of CM-III for these five attributes, although it does no harm to use the OMS in place of the RS despite the insignificant increment.

b) Optimal Monotonic Scores (Maximizing Predictability)

In contrast with our previous procedure for maximizing the interrelations among a set of variables (the homogeneity issue, whose primary application is in the area of aggregation of indices), our present concern deals with the problem of prediction. These two problems do not necessarily have the same solution, a matter which we shall address after presenting the CM-V method for obtaining OMS.

Given a set of n fixed independent variables, what is the best linear combination of those variables which will predict the responses to a dependent variable? This is, of course, the standard problem of linear regression analysis, i.e., we have the following linear model (when operating on standardized variables):

$$(6) \hat{y}_i = \sum_j \beta_j x_{ij},$$

where the x 's are our standardized set of independent variables, the β 's are the normalized beta weights, and the \hat{y} 's are the predicted scores ($i=1,2,\dots,N$ observations; $j=1,2,\dots,n$ independent variables).

Suppose, however, that the dependent variable, y , had undergone some unknown monotonic transformation. Under this condition we would obtain something less than monotonic transformation. Under this condition we would obtain something less than the best in terms of predictability. That is, if we could somehow recover that monotonic transformation, apply its inverse to make the regression system linear, then we would increase predictability. This, in essence, is the goal of CM-V OMS. We achieve this aim by iteratively computing: 1) the multiple correlation between the fixed independent variables and our current estimate of y^* , a monotone transformation of the dependent variable, and, 2) permuting the predicted scores into the order of y . When the correlation between predicted and monotone scores is a maximum, the process comes to a halt and we conclude by giving the prediction equation as well as the scale values for y^* .

In applying CM-V to data it is not necessary (and, indeed, may be contraindicated) that the set of x 's be homogeneous or unidimensional (although they may well be from a unified behavior domain relevant to the predictor). All that is required is that independent portions of the criteria be sampled by each of the x 's, for otherwise all of the predictability in the system will be carried by a subset of the x 's (as in a simplex structure, for example). A computationally simple example does not readily come to mind, so we shall illustrate CM-V by using our previous data on the five attribute satisfaction scales (OMS from CM-III) to predict rated satisfaction for luncheon meats. To the extent we are successful in predicting luncheon meat satisfaction, then one's satisfaction with luncheon meats can be considered to be a linear combination of one's satisfaction with the attributes of luncheon meats. If the multiple correlation is sufficiently high, then we have evidence that the appropriate attributes had been selected. On the other hand, if R (the multiple correlation) is too low, then either further attributes are needed to increase predictability or we have the wrong model.

We list the zero order correlations between the five OMS attributes and the RS ratings on luncheon meats satisfaction seriatim: .39 (packaging), .77 (taste), .64 (nutritional value), .36 (availability), and .47 (price). The corresponding beta weights are: .03, .60, .13, -.00, and .13, yielding $R = .785$. After a few CM-V iterations we achieved an R of .789, representing a modest increase of less than one per cent (.6%) in predictive variance. The zero order correlations between the five OMS attributes and the monotone transformation of the luncheon meats satisfaction scale (with their beta weights) are: .41 (.06), .77 (.60), .65 (.13), .38 (.02), and .47 (.11). Here again we find that the system is essentially linear and no substantial benefit accrues from scaling. Indeed, practically all of the predictive variance is covered by but the single attribute of taste ($r = .77$), which strongly suggests that we may well have a simplex structure (Guttman, 1954) in this instance.

One implication of the simplex model in the present case is that if we wish to improve our predictability of luncheon meats satisfaction, we should seek an attribute like taste, which is correlated with taste but has a higher

correlation with luncheon meats satisfaction. We will not dwell further on this issue, other than to point out that our exercise was not completely without merit, since it served to illustrate our methods and was suggestive of a more fruitful explanatory model. For other products, however, a more complex model has been found necessary. We shall conclude this section by giving the OMS for luncheon meats derived from CM-V (after adding in the mean of 2.98) and the scale interval increments, i.e.:

RS	1	2	3	4	5	6	7
y*	1.70	2.48	3.14	3.75	4.06	4.48	5.00
Δ		.78	.66	.61	.31	.42	.52

III. CONCLUSION

1. Two Models for Constructing Indices of Consumer Satisfaction

We shall now formalize our two approaches for constructing an index of consumer satisfaction (ICS) based on our discussion of the two methods (CM-III and CM-V) used for obtaining OMS. Although we shall confine our discussion to products and their attributes, the rationale is simply generalized to product classes and global satisfaction ratings.

The following notation will be helpful in discussing our two models:
Let

(7) $A = (a_{ijk})$ = the matrix of raw satisfaction ratings (RS) for the attributes of products ($i=1,2,\dots,N$ observations; $j=1,2,\dots,n$ attributes; $k=1,2,\dots,m$ products);

(8) $A^* = f(A)$ = a monotone transformation of the raw satisfaction scores for attributes, i.e., the OMS (via CM-III for each product separately);

(9) $P = (p_{ik})$ = the matrix of RS for products ($i=1,2,\dots,N$; $k=1,2,\dots,m$); and,

(10) $P^* = f(P)$ = a monotone transformation of product RS (via CM-V for each product, the dependent variable, and its associated scaled attributes, the independent variables, separately).

Now if, as a result of our analysis, we found that a particular product's satisfaction rating was predictable from one or at most two attributes (as we did find in respect to luncheon meats), we would include the product and its attributes in a CM-III analysis to yield the following measure of satisfaction for person i on product k :

$$(11) \quad s_{ik} = \frac{1}{n+1} \left(\sum_{j=1}^n a^*_{ijk} + p^*_{ik} \right),$$

where P^* now represents the OMS (via CM-III). The reader will easily recognize Eq. (11) as the simple aggregation model, whose rationale stems from a simplex structure.

If, on the other hand, we found that the structure was not that of a simplex, i.e., the network of interrelationships among the attributes was complex, then the appropriate model for obtaining subject i 's satisfaction score on product k would be:

$$(12) \quad s_{ik}^* = \sum_{j=1}^n \beta_j a_{ijk}^* + e_{ik},$$

or the linear-monotone regression model of CM-V, where s_{ik}^* represents the OMS and e_{ik} is the residual or error score.

In the first or aggregation model we will be in error to the extent that the set of attributes and product departs from a simplex structure. In the second or linear-monotone regression model we will incur error to the degree that the residual score component is large (i.e., R is low). In either event, we shall have achieved the best possible scores based on the data we have from the point of view of either homogeneity or predictability, respectively.

We defer a discussion of the details for constructing an ICS, since it simply involves an expansion of the above treatment for product classes and global ratings, summing over individuals (or subsets of individuals for group comparisons), and, possibly, weighting the normed index for a base period to take into account the economic importance of the various products.

In this paper we surveyed in general terms the contributions that traditional measures of economic performance can make to the measurement of economic and market performance. We noted, however, that these measures should be extended to take into account many phenomena which can be expressed in monetary and physical terms. Beyond that, social indicators are called for which monitor the quality of life and of social existence, including subjective welfare and satisfaction.

We discussed the methodological problems and opportunities arising from nonmetric scaling techniques, with special reference to our work on Indices of Consumer Satisfaction and Indices of Citizen Satisfaction. Dimensional simplification via monotone distance analysis, monotone vector analysis, and probability evaluated partitions techniques were described in general terms. Therefrom we went on to chart our approaches to scaling and aggregation of responses, including the uses of optimal monotone scores.

The methodology outlined in the latter part of this paper suggests a useful approach to supplementing objective indices. It offers a handle on the difficult problem of representing goods and services in terms perceived by the consumer and citizen. It takes these as inputs into a procedure for aggregating subjective satisfaction scores. These, in turn, lead to profiles of satisfaction or to an overall Index of Satisfaction.

Up to this point we have only investigated subjective welfare and satisfaction at one instant of time in respect to a very select set of items and for a small sample of individuals only. It would be important, both theoretically and practically, to extend our results for generality over time, items, and samples.

FOOTNOTES

1. An earlier version of this paper was presented as "Measurement of Subjective Welfare and Satisfaction" at the 84th Annual Meeting of the American Economic Association jointly with the Association for the Study of the Grants Economy, New Orleans, Dec. 27, 1971. The research was financed by Contract No. 12-17-05-1-610 from the U.S. Department of Agriculture.
2. Associate Professor of Psychology, University of Michigan.
3. Professor of Economics and Operations Research, Wayne State University; and Professor of Economics, University of Augsburg.
4. This is one of the major areas of concern for the Association for the Study of the Grants Economy which has studied not only monetary transfers between different groups but also the nature of non-market production. See for example, Pfaff & Pfaff, and Boulding.
5. For a bibliography on social indicators, see Knezo.
6. For a model of this type see Pfaff.
7. All methods in the Guttman-Lingoes series have been programmed for computers. For a description of these methods see Lingoes.

References

- [1] Arrow, K. Social Choice and Individual Values. New York: John Wiley & Sons, Inc., 1951.
- [2] Bain, J. S. Industrial Organization. (Second Edition). New York: John Wiley & Sons, Inc., 1968.
- [3] Baumol, W. J. Community Indifference. Review of Economic Studies, 1946-1947, 14.
- [4] Bergson, A. A Reformulation of Certain Aspects of Welfare Economics. Quarterly Journal of Economics, 1938, pp. 310-334.
- [5] Boulding, K. E. The Economics of Love and Fear: A Preface to Grants Economics. Belmont, Calif.: Wadsworth Publ. Co., 1972.
- [6] Bradburn, N. M. & Caplovitz, O. Reports on Happiness. Chicago, Ill.: Aldine, 1965.
- [7] Bradburn, N. M. The Structure of Psychological Well-being. Chicago, Ill.: Aldine, 1969.
- [8] Brayfield, A. H., & Crockett, W. H. Employee Attitudes and Employee Performance. Psychological Bulletin, 1955, 52, pp. 396-424.
- [9] Buzzell, R. D. Value Added by Industrial Distributors and Their Productivity. Columbus, Ohio: Ohio State University, 1959.
- [10] Caves, R. American Industry: Structure, Conduct, Performance. (Second Edition). Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1967.
- [11] Cantril, H. The Pattern of Human Concerns. New Brunswick, N. J.: Rutgers University Press, 1965.
- [12] Cattell, R. B. The Relational Simplex Theory of Equal Interval and Absolute Scaling. Acta Psychol., 1962, 20, pp. 139-158.
- [13] Cook, V., A Concept of Market Performance. Philadelphia, Pa.: Marketing Science Institute Working Paper P-51-2, May, 1968.
- [14] Coombs, C. A Theory of Data. John Wiley and Sons, Inc., New York, 1964, xviii+585.
- [15] Cox, R., & Goodman, C. S. The Marketing of Housebuilding Materials. Journal of Marketing, July, 1965.
- [16] Cox, R. & Goodman, C. S. Channels and Flows in the Marketing of Housebuilding Materials. Philadelphia, Pa. (Mimeographed), 3 Vols.

- [17] Douglas, S. A Review of Flow Analysis as a Measure of Marketing Performance. Philadelphia, Pa.: Marketing Science Institute Working Paper P-51-4, February, 1969.
- [18] Douglas, S. Marketing Performance: A Conceptual Framework. Philadelphia, Pa.: Marketing Science Institute Working Paper, P-51-1, June, 1968.
- [19] Drewnowski, I. (UNRISD). The Level of Welfare Index. Working Paper, January, 1968.
- [20] Easterlin, R. A. Does Economic Growth Improve the Human Lot?" (mineograph), 46 pages.
- [21] Green, P. E. & Carmone, F. J. Multidimensional Scaling and Related Techniques in Marketing Analysis. Boston, Allyn and Bacon, Inc., 1970, xv+203.
- [22] Gurin, G., Veroff, J. & Field, S. Americans View Their Mental Health. New York: Basic Books, 1960.
- [23] Guttman, L. A New Approach to Factor Analysis: The Radex. In: Lazarsfeld, P. F. (Ed.), Mathematical Thinking in the Social Sciences. The Free Press, Glencoe, Illinois, 1954, pp. 258-348.
- [24] Guttman, L. The Development of Nonmetric Space Analysis-A Letter to John Ross. Multiv. Behav. Res., 1967, 2, pp. 71-82.
- [25] Guttman, L. Measurement as Structural Theory. Psychometrika, 1971, 36, pp. 329-347.
- [26] Guttman, L. A General Nonmetric Technique for Finding the Smallest Coordinate Space for a Configuration of Points. Psychometrika, 1968, 33, pp. 469-506.
- [27] Hicks, J. R., Foundations of Welfare Economics. Economic Journal, 1939, 49, pp. 696-712.
- [28] Inkeles, A. Industrial Man: The Relation of Status to Experience, Perception, and Value. American Journal of Sociology, 1960, 66, pp. 31.
- [29] Kaldor, N. Welfare Propositions in Economics. Economic Journal, 1939, 49, pp. 549-552.
- [30] Knezo, G. I. The Social Sciences and Public Policy: A Selected, Annotated Bibliography. Washington, D. C.: The Library of Congress, HN US/70-162 SP, June, 1970.
- [31] Levin, C. H. Marketing Productivity Analysis. St. Louis: McGraw-Hill Publishing Co., 1965.

- [32] Lingo, J. C. An IBM 7090 Program for Guttman-Lingo Smallest Space Analysis-III. Behavioral Science, 1966, 11, pp. 75-76.
- [33] Lingo, J. C. An IBM 360/67 Program for Guttman-Lingo Conjoint Measurement-III. Behavioral Science, 1968, 13, pp. 421-422.
- [34] Lingo, J. C. Some Boundary Conditions for a Monotone Analysis of Symmetric Matrices. Psychometrika, 1971, 36, pp. 195-203.
- [35] Lingo, J. C. A General Nonparametric Model for Representing Objects and Attributes in a Joint Metric Space. In: Jardin, J.-C., (ed.), Archéologie et Calculateurs. C.N.R.S., Paris, 1970, pp. 277-298.
- [36] Lingo, J. C. A General Survey of the Guttman-Lingo Nonmetric Program Series. In: Shepard, R., Romney, A. K., & Nerlove, S., (Eds.), Multidimensional Scaling: Theory and Applications in the Behavioral Sciences. Seminar Press, Vol. I, 1972, pp. 49-68.
- [37] Lingo, J. C. A Fortran IV (G) Program for Guttman-Lingo Conjoint Measurement-V. Behavioral Science, 1972, 17, pp. 333-334.
- [38] Lingo, J. C. The Guttman-Lingo Nonmetric Program Series, 1972, (in press).
- [39] Lingo, J. C. & Cooper, T. PEP-I: A Fortran IV (G) Program for Guttman-Lingo Nonmetric Probability Clustering. Behavioral Science, 1971, 16, pp. 259-261.
- [40] Lingo, J. C. & Guttman, L. Nonmetric Factor Analysis: A Rank Reducing Alternative to Linear Factor Analysis. Multiv. Behav. Res., 1967, 2, pp. 485-505.
- [41] Lingo, J. C. & Roskam, E. A Mathematical and Empirical Study of Two Multidimensional Scaling Algorithms. Michigan Mathematical Psychology Program, 1971, 1, pp. 1-169.
- [42] Pfaff, M. Theories of Market Systems: Implications for the Measurement of Market Performance. Philadelphia, Pa.: Marketing Science Institute Working Paper, April, 1968.
- [43] Pfaff, M. The Marketing Function and Economic Development: An Approach to a Systemic Decision Model, Philadelphia, Pa.: Marketing Science Institute, 1968.
- [44] Pfaff, A. B. Productivity and Value Added Measures in Marketing. Philadelphia, Pa.: Marketing Science Institute Working Paper, P-51-9, November, 1969.
- [45] Pfaff, A. B., & Pfaff, M. Methods of Welfare Economics in the Measurement of Market Performance. Philadelphia, Pa.: Marketing Science Institute Working Paper P-51-7, May, 1969.

- [47] Pfaff, M. & Pfaff, A. With an Introduction by K. E. Boulding, The Grants Economy. Belmont, California: Wadsworth Publishing Co., 1972.
- [48] Reder, M. Studies in the Theory of Welfare Economics. New York: Columbia University Press, 1947.
- [49] Robbins, L. An Essay on the Nature and Significance of Economic Science. London: Macmillan, 1932.
- [50] Roskam, E. Metric Analysis of Ordinal Data in Psychology. Voorschoten, Nijmegen, Holland, 1968, pp. 1-165.
- [51] Roskam, E. & Lingoos, J. C. MINISSA-I: A Fortran IV (G) Program the the Smallest Space Analysis of Square Symmetric Matrices. Behavioral Science, 1970, 15, pp. 204-205.
- [52] Rothenberg, J. The Measurement of Social Welfare. Englewood Cliffs, N. J.: Prentice-Hall, 1961.
- [53] Ruggles, N. & Ruggles, R. The Design of Economic Accounts. New York: National Bureau of Economic Research, 1970.
- [54] Smith, P. C. & Cranny, C. J. Psychology of Men at Work. Annual Review of Psychology, 1968, 19, pp. 467-496.
- [55] Smith, P. C. Kendall, K. M. and Hulin, C. L. The Measurement of Satisfaction in Work and Retirement: A Strategy for the Study of Attitudes. Chicago, Ill.: Rand McNally & Co., 1969.
- [56] Stagner, R. Perceptions, Aspirations, Frustrations, and Satisfactions: An Approach to Urban Indicators. The Annals of the American Academy of Political and Social Science, Philadelphia, 1970, 388, pp. 59-68.
- [57] Stern, L. W. Market Structure Analysis in the Measurement of Market Performance. Cambridge, Mass.: Marketing Science Institute Working Paper P-51-6, May, 1969.
- [58] United Nations Research Institute of Social Development (UNRISD). Social and Economic Factors in Development. Report No. 3, February, 1966.
- [59] UNRISD. The Level of Living Index. Report No. 4, September, 1966.
- [60] UNRISD. Measuring Social Variables in Real Terms. Working Paper, February, 1968.
- [61] UNRISD. The Level of Living Index: New Version. Working Paper, March, 1968.
- [62] Wilson, W. Correlates of Avowed Happiness. Psychological Bulletin, 1967, 67, pp. 294-306.

AN INDEX OF CONSUMER SATISFACTION¹

Anita B. Pfaff²
Wayne State University,
and University of Augsburg

I. Consumer Satisfaction in a Market Economy

The ultimate claim and justification of a market economy is its ability to satisfy individual consumers better than any other type of economic system. In the face of differing ideologies and organizational forms competing for man's allegiance, this is no small claim and no mean achievement. Yet, in a period when product markets have vastly expanded the offering of goods, gadgets, and services, and when factor markets have provided for increased affluence to the majority of Americans, there has arisen a consumer restlessness and even outright revolt against the system. Indeed, the voices of criticism have come not only from our urban ghettos or rural enclaves which have been bypassed by this economic upswell like islands of misery in an ocean of affluence. The suburban housewife, favored most by the spectacular increase in living standards of the 1960's, sets the tone in the growing chorus of protest. No doubt, general economic forces associated with inflation and unemployment have made the consumer pessimistic and more aware of the limits of the household budget. But complaints pertain often to specific facets of product or service offerings and to specific parts of the whole distribution process. Even more pronounced, seemingly, is the voice of criticism levied at the quality of services offered in the market, ranging from poor automobile servicing and repairs, to more sophisticated services offered by professionals.

This is a reaction not only against the general economic scene, or the institutions and individuals constituting the retail levels, but at the whole mass-marketing process.

Marketers generally take great pride in their strategies of product differentiation which promise a plethora of product types to satisfy every whim of the consumer, as long as it is backed up with the necessary purchasing power. Strategies of market segmentation, in turn, are geared to tailoring product and service offerings to the desires of specific consumer groups. In the face of all this one may be somewhat surprised at the discrepancy that exists between the images of suppliers and the images of consumers about the performance of products and markets.

The aim of the research project reported in this paper is not to level further criticism at the "goose that lays the golden eggs." Rather, it attempts to marshall modern psychometric techniques--particularly nonmetric scaling techniques and related breakthroughs in computational technology--to the task of formulating social indicators of market performance. These indicators may serve the ends of business and industry aiming at improving their own performance, and of government charged with the task of providing happiness to its citizens. But we hope that such information will ultimately benefit

consumers who will be the beneficiaries of improved market performance. An index of consumer satisfaction will not replace marketing research, but rather indicate a point on departure for such efforts.

II. Measures of Economic Performance

1. Objective Measures

Over the last four decades, economists have become increasingly concerned with the evaluation of economic performance. Such evaluation is contingent on two prerequisites: First, the goals of the economic system have to be known; and second, measurements on variables representing these goals, or operational "proxies" of these goals, have to be carried out, to determine to what degree these goals have been achieved.

As a first step in this direction, a number of familiar economic indicators have been developed--such as gross national product and related concepts on the output side, or labor force participation and investment rates on the input side. Consumers' living costs are estimated with the aid of the Consumer Price Index. These economic indicators monitor economic processes at regular intervals. No doubt, these indicators provide valuable information about how well our economy performs. They suffer, however, from a monetary and physical bias, as all effects have to be expressed in monetary or physical form. They fail in helping answer questions such as: "Is enough produced to provide everybody with the most essential goods and services? Are the goods and services produced that people want most? Are people satisfied with market offerings?"

Objective measures of economic performance can act as reasonable proxies of subjective welfare only if certain questions are excluded from the menu of economic topics: To do this we have to work with certain economic ideal types: Monetary and physical measures of performance will reflect subjective welfare provided, first, we assume the distribution of income to be given and outside the area of concern of the economist; second, the price mechanism operates in a perfectly competitive market; and it guarantees that relative prices reflect the demand patterns of the consumers. The demand function, moreover, is an expression of "effective" demand, i.e., willingness backed by ability to pay for the goods and services. With a given income distribution it is quite conceivable that a price structure results which provides essentials to the poor in far too small amounts, and luxuries to the rich in relative abundance. An increase in the output of luxuries may result in an upward movement of the objective economic indicators, suggesting an improvement in social welfare; while by many social norms, and in the perception of most people, the change would not necessarily be considered desirable.

In a perfectly competitive market the price mechanism will ensure that the right products will be produced in right amounts. Imperfections in the market are however, the rule rather than exception. This means that the "feedback action" of the market is considerably weakened. The prevailing price structure is influenced by differential power, special skills in negotiation, and a variety of factors extraneous to true demand. While the objective economic indicators are thus measurable at least in principle, they may not satisfy

the first requirement of an evaluative theory of the socio-economy, since they fail to measure what the actual goals are.

Considerable attention has recently been given to national goals and the development of social indicators. Increasingly these studies do not only concern themselves with physical or objective measures, but with subjective measures as well.

2. The Index of Consumer Satisfaction: A Measure of Subjective Welfare

The paper reports on a study designed to develop a "subjective" indicator of consumer satisfaction, as a complement to the Consumer Price Index (CPI). The CPI measures variations in the prices of a general "market basket." The Index of Consumer Satisfaction (ICS) is to supplement this "cost measure" with a "benefit measure." It tries to determine the benefits that consumers receive--or the degree of satisfaction they experience--from the operation of the market. Since no convenient aggregate measure of utilities exists that provides an easy and reasonably unambiguous measure of satisfaction, consumers' opinions have to be explored on the subject. An index, or a profile of indices, has to be based on their avowed satisfaction. In the absence of weight changes which express the quantitative composition of the "market basket," the CPI is of the form:

$$(1) \text{CPI}_{i:o} = \frac{\sum_j G_{oj} P_{ij}}{\sum_j G_{oj} P_{oj}} \times 100$$

where G_{oj} is the weight of the j^{th} product in the market basket in the base period, and P_{ij} and P_{oj} are the average prices of the j^{th} product at time i and the base period, respectively.³

An analogous Index of Consumer Satisfaction of the form

$$(2) \text{ICS}_{i:o} = \frac{\sum_j W_{oj} S_{ij}}{\sum_j W_{oj} S_{oj}} \times 100$$

may be suggested where W_{oj} represents the weight and S_{ij} the satisfaction scores of the j^{th} product, in current and base periods, respectively. The weights W_{oj} and W_{ij} would not generally coincide with G_{oj} and G_{ij} . The latter relate to average amounts purchased. The former constitute "importance weights"--the weights consumers associate subjectively with given goods or services. Some items tend to be purchased in large quantities. Yet, they may not rank very highly in terms of subjective importance in the eyes of the consumer.

Similar to the G's in the Consumer Price Index, the W's would be revised only occasionally, to reflect a change in attitudes or distribution of income and wealth. They are likely to differ between different socioeconomic groups.

There is no prima facie basis for assuming additivity of satisfaction scores across products. Yet in a situation of changing relative prices, the simple additivity assumption underlying the CPI may be questioned for precisely the same reason: This is due to the interdependence of products with regard to substitutability or complementarity. The additivity assumption is no less heroic when applied to the CPI than to the ICS.

III. The Design of the Pilot Study

A pilot study was conducted in order to explore whether the approach to measurement of consumer satisfaction was technically and economically feasible and whether it yielded results which have interesting policy implications.

The sampling procedure, and the conduct of the field surveys was carried out by the Survey Research Center of the Institute for Social Research of the University of Michigan. A sample was taken in spring of 1971, consisting of 574 family units with both husband and wife present, and with small children no older than ten. This random sample was drawn on the basis of birth records and other public records. Subsequent cross-checks indicated that there were essentially similar distributions among the demographic descriptors as in the 1971 Survey of Consumer Finances. A small income bias was noted in our sample due to a slight under-representation of low income families.

A limited number of products were chosen for inclusion in the pilot questionnaire. A duplication of the entire CPI market basket would have resulted in much too long a questionnaire. The development of an index based on a more extensive product or service selection would have to follow an extension of the common CPI practice. This entails the collection of responses on different items from different but similar subsamples.

It does not suffice to determine only the consumers' overall satisfaction with a product. A consumer acquires a bundle of goods and service combinations in any purchase. Various marketing aspects--packaging, display, etc.--represent relevant dimensions of the product. The product itself may exhibit a number of attributes. For example, a housewife who purchases a refrigerator may consider not only the objective or engineering-economic attributes, such as size, internal arrangement, economy of operation, or special features such as automatic ice maker. She also looks at marketing attributes such as price, availability, convenience of store location, availability of repair services, image, and so on. Finally, a model is likely chosen which is perceived as particularly satisfactory with regard to some of these attributes, and not so satisfactory with regard to others. The rational decision-maker will choose a product which conforms more to his desires in those attributes that he perceives as important to him. His overall satisfaction with the product may thus be seen as a composite of his satisfaction with the various attributes.

The more salient aspects are likely to dominate the picture. Random fluctuations in the product image are likely to occur as a consequence of particular experiences involving the product. But over time, a fairly stable image is likely to emerge. The stability of the product image will also depend on the frequency of purchase and, therefore, the product life.

For purposes of this analysis a very limited number of different products was included. The study was directed at the investigation of products with different purchase frequencies and life. The overall sample was divided into two subsamples--A and B. In subsample A, husbands were interviewed about their overall satisfaction with the house they lived in and their degrees of satisfaction with various attributes of the house. In subsample B, husbands responded to questions about their car and its attributes. In subsample A, wives answered questions about food products; and in subsample B, wives were asked about clothing. In both subsamples the wives were also interviewed about their general satisfaction with all products they purchased, with breakfast cereals, and five attributes of breakfast cereals, luncheon meats, and five attributes of luncheon meats, as well as women's clothing, such as dresses, blouses, slacks, shirts. They were also asked about their satisfaction with food in general, and appliances.

Satisfaction was measured on a seven-point scale ranging from "very satisfied" (A) to "not at all satisfied" (G). Respondents were asked how important each of the attributes were for the purchasing decision. Importance was measured on a seven-point scale as well, ranging from "very important" (A) to "not at all important" (G). Letters rather than numbers were associated with the scale in order not to suggest a particular order or specific quantitative relation between points on the scales. A "satisfaction" scale (rather than a scale ranging from very satisfied to very dissatisfied) was employed: It was felt that a scale of the latter type is a mixed scale, derived from two separate scales where one measures satisfaction and the other dissatisfaction. These two scales are monotonically related, but they may not coincide.

The choice of products was based on the different purchase frequencies and durability (or span of life). "House" was chosen as a product with very long life span, characterized by infrequent purchasing (or renting) decision. Cars occupy an intermediate position. They are shorter-lived than houses. Yet, they are not purchased very frequently. Moreover, the high expense associated with their purchase makes the decision also important in terms of the commitment of the family budget. Clothing occupies the next position on a life span scale. Finally, luncheon meats and cereals are purchased frequently and have a very short life span.

The overall sample was reduced to eliminate all individuals and their spouses who had responses on the 47 satisfaction scales (subsequently called "internal" variables) such as "do not know," "not applicable," "inappropriate," "uncertain." To be included in the analysis, both husband and wife had to respond to all the internal satisfaction questions put to them on the 7-point scale. This reduced the sample to size 342 men and women with subsample A consisting of 189 and subsample B of 153 couples.

Complete ratings on a 7-point scale regarding 37 questions about importance of product attributes were also obtained for the reduced sample of 342 dyads.

IV. Results of the Analysis of Satisfaction Scores

Both raw scores and optimal weighted scores were employed for measuring consumer satisfaction.

1. Raw Scores

The first step, consisting of a metric measurement of satisfaction, was taken by associating numerical values with alphabetic response. Arbitrarily, numbers 1 through 7 were assigned to the letters A (very satisfied) through G (not at all satisfied). These are subsequently called "raw scores." This particular scale was used for all individuals. It implies that respondents perceived adjoining points as "equidistant." In other words, individuals perceived a movement from 1 to 2 as psychologically equally easy--or difficult--as a movement from 5 to 6. It also implies that no semantic differences with respect to the two end points (or any other point with verbal association) exist.

2. Optimal Weighted Scores

A short digression is needed to show why raw scores appeared not very useful. On a priori grounds two extreme hypotheses on the nature of satisfaction may be considered. One may hypothesize that individual products and attributes are independent of each other in terms of the satisfaction they yield. Alternatively, one may postulate that satisfaction generalizes: Persons who are more satisfied with one product are also more likely satisfied with another. If the former hypothesis were true, the satisfaction created by each product would be an independent random variable; and an Index of Consumer Satisfaction would be based on the sum or a mean derived from the set of products and individuals. In an aggregation across products, information would obviously be lost, since one index number may result from a large number of score distributions. If the second hypothesis were to be true, this would be reflected in high correlations between satisfaction scores of attributes and products, suggesting a generalization of satisfaction. In the latter case an additive or averaging aggregation across products would not entail too great an information loss. It would also have important policy consequences. We did not perform a rigorous statistical test on the significance of the simple linear correlation coefficients of satisfaction ratings based on raw scores. Simple inspection sufficed to reject the hypothesis of a generalization of consumer satisfaction. Most correlations had a positive sign; but most of them were also near zero. Due to the restrictive assumptions of the raw scores, such a result may well have been statistical artifact.

To investigate this possibility, a set of optimal scores was derived by application of Multivariate Analysis of Contingencies III (MAC III).⁴ This involved a rescaling of the raw scores in such a fashion as to maximize the (new) coefficients of correlation. Thus a maximum degree of homogeneity (= average correlation) will be achieved.

how satisfied are you with	RAW SCORES					η^2
	very satisfied		not satisfied			
	1	2	3	4	5	
1. your standard of living	-95	45	-15	-61	-60	.040
2. your standard of living, if it were the same 5 yrs. from now	146	72	-17	0	-39	.041
3. breakfast cereal	63	6	-27	-62	-98	.077
4. its nutritional value	92	30	-18	-42	-59	.086
5. its taste	61	-15	-52	-83	-71	.087
6. its packaging	67	-18	-32	-72	-67	.072
7. its availability	42	-45	-75	-96	-100	.061
8. its price	81	68	28	3	-60	.072
9. luncheon meat	63	43	-19	-52	78	.111
10. its packaging	70	9	-37	-69	-64	.085
11. its taste	67	17	-43	-69	-65	.119
12. its nutritional value	80	20	-2	-60	-73	.113
13. its availability	65	-22	-56	-55	-64	.096
14. its price	116	53	58	-9	-49	.109
15. clothing (women's)	87	45	-5	-38	-71	.126
16. its wearability	81	35	6	-38	-72	.128
17. its style	69	14	-39	-70	-81	.088
18. ease of upkeep	58	-10	-45	-100	-80	.082
19. comfort	60	-19	-60	-87	-84	.102
20. service	89	36	0	-31	-76	.076
21. price	117	70	44	5	-49	.104
22. credit terms	56	79	21	-30	-56	.055
23. food	95	39	-12	-39	-103	.074
24. clothing	101	44	-12	-59	-79	.131
25. appliances	69	-31	-65	-37	-15	.060

Table 1: Optimal Weighted Scores (based on MAC III for 25 Satisfaction Scales, 5 Categories, Female Respondents).

MAC III does not guarantee a monotonic rescaling. However, in the absence of a more complex theory of interdependence of product and attribute satisfaction involving non-linear relations, a non-monotonic rescaling of raw scores appears void of meaning. In other words, it may be difficult to interpret scores on a scale ranging from "very satisfied" to "not at all satisfied" (such as, -101, 12, -23, 8, 61, 20, 22). On the other hand the imposition of a monotonicity constraint on the new scores results in a lower degree of homogeneity. The degree of homogeneity or average intercorrelation achieved by MAC III scores depends on sample size, number of categories, and number of individuals falling in each category.

The original seven categories proved inconvenient: A rather small number of individuals fell into raw-score categories 6 and 7 (very low satisfaction scores). This tended to distort the scales; and it led to non-monotonic scales. Categories 5, 6, and 7 were therefore collapsed into one category. The 5-point scale proved quite appropriate.

Of the 47 satisfaction scales 28 are perfectly monotonic and the others show only minor deviations. The degree of homogeneity could be improved; but the average intercorrelations are still not very high.

Table 1 shows the optimal MAC III scores for a system of 25 satisfaction responses of females on (1) standard of living, (2) breakfast cereals plus five attributes, (3) luncheon meat plus five attributes, (4) women's clothing plus seven attributes, (5) food, (6) clothing in general, and (7) appliances. Inspection of the right most column (η^2) shows that the system is not homogeneous. The highest average correlation (η) can be observed for clothing in general; even this amount is less than .4 ($\eta^2 = .131$).

A much higher degree of homogeneity can be achieved by analyzing a smaller system of more closely related variables. Table 2 shows, for example, the results of the analysis of satisfaction with luncheon meat and its attributes. Higher values of η^2 can be noted therefrom.

How satisfied are you with	Raw Scores					η^2
	1	2	3	4	5	
1. luncheon meat	62	40	-10	-46	-86	.326
2. its packaging	63	25	-41	-64	-77	.192
3. its taste	62	22	-33	-59	-85	.356
4. its nutritional value	73	27	1	-55	-82	.307
5. its availability	61	-9	-46	-79	-74	.179
6. its price	109	59	55	-4	-51	.206

Table 2: MAC III Scores of Satisfaction with Luncheon Meat

(5 attributes; female respondents).

However, even in this system the highest average correlation r is less .6 ("taste"). With the exception of "availability" all variables have strictly monotonic scores.

All MAC III scores are generated in such a fashion as to have a mean of zero and the equal variance. This has a distinct advantage in comparing relative well-being when absolute comparisons between scales are difficult. For example, being "very satisfied with the price of luncheon meat" is indicated by a score of 109, whereas being "very satisfied with packaging" is denoted by a score of 63. The higher value results from a relatively small number of respondents in the first category, and from the relatively low mean satisfaction with price of luncheon meat (4.39 in terms of raw scores, as compared to a mean raw score of satisfaction with packaging of 2.63). The person who is "very satisfied" with "price" generally may be better off than the person who is "very satisfied with packaging." The higher score of 109 for "satisfaction with price" would thus be justified. MAC III scales of the first four variables are rather similar. The "availability" scale reflects a rather high level of satisfaction. This is evident from the large number of negative signs; a response in the second category in terms of this population means relatively low satisfaction. This scale is slightly nonlinear: Category 5 has a higher score than category 4.

Table 3 exhibits MAC III scores of male respondents for more general satisfactions. The four variables under analysis are responses to the following questions: (1) "Generally speaking, how satisfied are you with the quality of all the products you buy?"; (2) "Compared to 5 years ago, would you say the quality of all the products you and your family buy is better now, worse, or about the same?"; (3) "What about the quality of service you get in the stores where you and your husband shop, is it better now than five years ago, about the same, or worse?"; (4) "What about getting things repaired properly; is it easier to have things repaired now than it was five years ago, harder, or about the same?"--While variable 1 was measured on a 7-point scale (collapsed to 5 categories), variables 2-4 contained an implicit satisfaction rating; it assumes that the response "better now" or "easier now" corresponded to very satisfied (category 1), "same" to category 3, and "worse now" or "harder now" to category 5. Obviously no MAC scores were derived for categories 2 and 4, since no responses fall into these categories.

Variable	Raw Scores					n^2
	1	2	3	4	5	
1. all products	39	55	-10	-74	-112	.061
2. quality change	79		6		-69	.083
3. service	99		25		-55	.119
4. repairs	91		48		-48	.108

Table 3: Optimal Weighted Scores (MAC III) for General Satisfaction, 342 Male Respondents.

The low level of η^2 indicates a high degree of heterogeneity in the system. Scale (1), moreover, exhibits slight nonlinearity. Concentration of responses in variables 2-4 were on "same" and "worse"; the fact that the score for the "same" category is above average (0) reflects this phenomenon.

Optimal weighted scores for importance responses of product attributes were derived. Since attributes had been selected for inclusion in the questionnaire on the basis of their salience, there occurred very few responses in categories 5-7 (not important). Moreover, the mode of all importance scores was in category 1 ("very important"). MAC III scores (not shown in this paper) proved not very useful, since most scales were non-monotonic with the raw scores.

3. Optimal Monotonic Scores

The definite disadvantage accompanying the use of optimal weighted (MAC III) scores is their potential lack of monotonicity to the implied order of the raw scores. In order to avoid this, a rescaling of raw scores has to incorporate the constraint of maintaining the original order.

The Conjoint Measurement III (CM III) algorithm (Lingoes, 1968) provides for such a rescaling: While original order is retained the homogeneity (or average linear correlation) between variables is maximized. Generally, the degree of homogeneity achieved by the use of CM III Scores will be less than that of the MAC III scores. In particular if large digressions from the original order occur in MAC III scores the retention of the original order will involve an opportunity loss in homogeneity.

The experience with MAC III scales, viz. the low degree of homogeneity achieved by combining rather diverse goods or concept suggested that satisfaction does not seem to generalize. Scales were, therefore, jointly developed for closely related sets of measures such as a product and its attributes.

Like MAC III scores, CM III scores depend on sample size and frequency distribution. The attribute variable scores are, however, standardized. This implies that the mean equals zero; and the standard deviation equals one. Since monotonicity is no problem, the original seven categories were retained for the raw score input. The product variable is rescaled to a zero mean as well.

The product "luncheon meat" and five attributes (packaging, taste, nutritional value, availability, price) were analyzed accordingly.

Table 4 summarizes some of the results of this analysis. The left hand side of the table shows the CM III scores for "luncheon meat" and 5 attributes, the scores run from negative figures ("very satisfied") to positive figures ("very dissatisfied").

RAW SCORES

How satisfied are you with	1	2	3	4	5	6	7	r_{yx_j}	$r_{y^*x_j^*}$
1. luncheon meat	-.867	-.386	.077	.497	.776	1.049	1.608		
2. packaging	-1.000	-.384	.230	.845	1.460	2.075	2.690	.3773	.3847
3. taste	-.988	-.397	.194	.784	1.375	1.966	2.557	.7830	.7784
4. nutritional value	-1.176	-.580	.017	.614	1.211	1.808	2.405	.6588	.6568
5. availability	-.858	-.206	.446	1.099	1.751	2.403	3.055	.3518	.3585
6. price	-1.759	-1.240	-.722	-.203	.315	.834	1.353	.4778	.4774
Average								.5297	.5306

Table 4: Optimal Monotone (CM III) Satisfaction Scores of Luncheon Meat and 5 Attributes.

Inversion of the signs would result in a "satisfaction" rather than the present "dissatisfaction" scale. The larger number of positive values, and their higher absolute values, indicate a positively skewed distribution: Most consumers reported to be fairly satisfied. The attribute "price," for example, shows a higher concentration in negative scores; this indicated that fewer people were "very satisfied" with price. The mean raw score of satisfaction with luncheon meat was 2.49, while the mean raw score of satisfaction with price of luncheon meat was 4.05. This denotes a higher level of dissatisfaction. Since the means of all scores are zero, these scores are useful in investigating relative distributions of satisfaction, rather than absolute levels.

Generally, CM scores of adjacent categories will be closer if a relatively large number of respondents fall into these categories. If only few respondents fall into an extreme-value category, the extreme values will differ by a large amount from the next value--and the mean: The scale has to discriminate between rather few consumers and the rest of the consumer group. This implies that a move to extreme values (or from extreme values) may be much harder than between other levels. The interval between adjoining categories can, therefore, be interpreted as a measure of the perceived distance between these response levels. With raw scores all intervals were of the same size. This may be an inadequate reflection of the true perception of the meaning of different degrees of satisfaction.

The right-hand side of Table 4 shows the correlation coefficients between the product variable measured in raw scores (y) and each attribute variable measured in raw scores (x_j , $j=1, \dots, 5$), in the column r_{yx_j} . The average of these coefficients (.5297) is reported in the bottom line ("average"). The column headed by r_{y*x_j} reports the corresponding correlation between satisfaction measured in terms of CM III scores. While some correlations are smaller than corresponding raw score correlations, an overall improvement in homogeneity is reflected in the somewhat higher average of .5306.

Both the small increase in homogeneity as well as the fact that the new scales almost exclusively exhibit little unequal stretching or shrinking of intervals suggest that in the instance of luncheon meat CM III scores, in fact, did not provide a great improvement in homogeneity. The almost equal intervals do not provide a great increase in sensitivity.

4. A Comparison of Scoring Systems

Raw scores have the advantage of allowing absolute comparisons between attributes or products. However, optimal weighted scores and optimal monotone scores provide higher homogeneity of measures and higher sensitivity to extreme responses.

It is questionable whether reporting being "very satisfied" (raw score 1) with a car and with a breakfast cereal rally means the same thing, and whether they should therefore be denoted by the same score. MAC III or CM III will generally render different scores for the two scales. This results from the different frequency distributions for the two products. Both will yield a mean score of zero and the same standard deviation for all scales. This

implies that the mean satisfaction with luncheon meat (2.98 in terms of raw scores) will be identically zero, whether measured in MAC III or CM III scores.

A satisfaction scale with an anchored mean may, however, be very useful for policy making. At present we have no generally accepted satisfaction scale such as the Fahrenheit scale for temperature, or a Murk Index for air pollution. Therefore, we can measure the present state of satisfaction and associate with it an average score of zero. Future deviations will be reflective in positive or negative averages. Similarly one can make comparisons at a given point in time between different subpopulations. It could be argued that the same may be achieved by using raw scores, by measuring their deviation from the mean. Such a measure recommends itself on the grounds of simplicity. But the drawback of such a system is the lower homogeneity of scales. If we were to contemplate a reduction in the number of products or attributes to be included in a survey, then we would have to choose one or a few out of a group of very homogeneous products or attributes.

5. Prediction and Aggregation of Satisfaction Across Attributes

"Mixed feelings" about a product are not uncommon. While some attribute appears uncommonly pleasing, another one may delight us not at all. An overall reaction to the item can still be elicited, which in all probability reflects and aggregates these "mixed feelings."

The simplest approach to arrive at an aggregate satisfaction measure would be to sum satisfaction scores across all attributes. Or, in order to facilitate comparison between different products with a different number of attributes, an average across attributes may be more appropriate. One might even include the general satisfaction score in the averaging process.

Under the assumption of a simple additive relationship the average ought to be close to the general product satisfaction score for any one individual. The simple averaging process assumes approximately equal importance of all attributes. Any scoring system may be used to derive such an average.

In Table 5, aggregate scores of 3, (out of 342) female respondents are shown for luncheon meat. They are based on satisfaction scores for the product as well as all attributes. The numbers in parentheses are the overall product satisfaction response scores. On the right hand side of the table the raw score responses are shown for luncheon meat and its attributes for the three respondents. As can be seen the attribute scores are either equal or below the product scores (underlined) which results in averages that indicate lower satisfaction than the general response does. This is, however, only typical for these three respondents. Particularly the second respondent seemed inconsistent, in reporting to be very satisfied with the products, but far less so with each and every of the attributes.

A sample of average satisfaction scores with three products (luncheon meat, breakfast cereals, women's clothing as well as average across satisfaction with "all products," "all foods," "all clothes," and "all appliances," and the average across the previous four averages are depicted in Table 6.

Satisfaction is measured in terms of MAC III scores. The mean of each set of aggregate scores across the entire sample is zero.

Individual	Raw	MAC III	CM II	Raw Score Profile
1	3.67 (3)	-30 (-10)	.37 (.08)	<u>332337</u>
2	3.17 (1)	-16.33 (62)	.14 (-.87)	<u>123436</u>
3	2.67 (2)	6.33 (40)	-.15 (-.39)	<u>222334</u>

Table 5: Average Satisfaction With Luncheon Meat Measured in Three Different Scoring Systems, for 3 Respondents.

From the very few examples in Table 5 we saw that average satisfaction sometimes differs very much from stated overall satisfaction; this was the case even though, as a correcting factor, the reported product satisfaction score was included in the averaging process. Had this not been the case, deviations would even have been larger. In other words, a simple average does not necessarily appear to be the best predictor of overall product satisfaction. The cause for this shortcoming may be in the differential impact that different attributes may have on the evaluation of the product as a whole. One way of recognizing this differential impact may be the computation of a weighted average.

An appropriate way of inferring implicit importance weights for different attributes consists in estimating regression coefficients for a model explaining variations in overall product satisfaction through variations in attribute satisfaction. Satisfaction scores in a regression model may again be measured in any one of the three scoring systems discussed above.

Regressions of product satisfaction on attribute satisfactions were performed for luncheon meat using all three scoring systems (raw, MAC III and CM III scores). The summary results are shown in Table 7.

Regression coefficients are reported in columns "coef.", their standard errors in columns headed by " \hat{s}_b ." One asterisk in columns "sig." indicates that the coefficient is significant at the 5 percent level, two asterisks, that the coefficient is significant at the 1 percent level. R denotes the multiple correlation coefficients, \hat{s}_{yx} the standard errors of the estimate, and "sequence" the sequence in which variables were entered as explanators in the stepwise multiple regression, which also coincides with the rank-order of the standardized regression coefficients. The poorest predictive power is exhibited by raw scores as a consequence of the lower homogeneity of the system. MAC III

Individual	Breakfast Cereal	Luncheon Meat	Clothing	All Products	Sum
1	-21.50	-29.50	-25.75	-77.50	-34.25
2	-48.00	-16.00	4.25	87.25	-3.87
3.	-1.00	6.17	31.25	-7.75	10.42
4	-9.00	43.83	17.87	8.50	16.08
5	16.83	5.83	-67.25	-13.50	-19.00
6	65.67	52.17	-23.63	-8.00	20.25
7	65.67	61.83	53.00	40.50	56.29
8	11.17	-27.00	-39.75	-6.25	-18.25
9	24.17	16.00	0.75	18.50	12.54
10	-15.00	17.00	20.38	24.00	11.29

Table 6: Average Satisfaction Score Profiles of 10 Respondents.

scores are better predictors than CM III scores. However, MAC III scores are not fully comparable with the others since they were based on 5 categories only as compared to 7 in other scores. It should be noted, that a higher degree of homogeneity does not necessarily entail a more accurate prediction.

	Raw Scores			MAC III Scores			CM III Scores		
	Coef.	S _b	Sig.	Coef.	S _b	Sig.	Coef.	S _b	Sig.
Constant	.003			-.138			.000		
1. packaging	.343	.039	**	.072	.040		.024	.026	
2. taste	.339	.057	**	.583	.052	**	.395	.034	**
3. nutritional value	.155	.043	**	.119	.050	*	.055	.035	
4. availability	.054	.055		.005	.039		.066	.025	**
6. price	.044	.034		.022	.040	**	.081	.027	**
R	.712			.789			.760		
S _{yx}	1.094			33.175			.438		
Sequence	1,2,3,5,4			2,5,3,1,4			2,5,4,3,1		
* significant at the 5% level									
** significant at the 1% level									

Table 7: Summary Result of Regression of Product Satisfaction on Attribute Satisfaction for Luncheon Meat.

The numerical impact of "taste" appears the largest in all scoring systems.

"Taste" and "price" are significant for CM III and MAC III scores; "availability" appears significant in CM III scores, and "nutritional value" in MAC III Scores. In raw score measurements "packaging," "taste" and "nutritional value" appear significant.

In the absence of a more convincing theory we assume that product satisfaction depends in a linear additive fashion on attribute satisfaction. If this is the case, regression results also provide a possible clue as to which attribute is the most instrumental determining overall satisfaction with the product. It may also provide suggestions as to which attributes have to be given a different "image" so as to change the consumer's attitude. This aspect is important in a policy making context. In this case regression is not only used as a prediction technique but also for the purpose of control.

This technique is not only useful in explaining product satisfaction in terms of attribute satisfaction. It can analogously be employed to determine the weight of individual products in the aggregate satisfaction of product groups. A rough attempt towards this end was made by trying to explain satisfaction with the quality of all products in terms of satisfaction with breakfast cereals, luncheon meats, women's clothing, food, all clothing and appliances. Obviously the choice of independent variables is not satisfactory, as other products influence overall satisfaction which were not included in the pilot survey.

The regression results are reported in Table 10.

	Coef.	S_b	Sig.
Constant	1.083		
1. breakfast cereal	.060	.035	
2. luncheon meat	.077	.032	*
3. women's clothes	.102	.042	*
4. food	.099	.046	*
5. all clothes	.233	.057	**
6. appliances	.065	.044	
7. R	.555		
8. \hat{S}_{yx}	.963		
Sequence	5, 2, 3, 4, 1, 6		

Table 8: Summary Result of Regression of Satisfaction With All Products of 6 Product Groups.

(*significant at the 5% level; **significant at the 1% level)

The regression was based on raw scores. The predictive accuracy reflected in a coefficient of multiple correlation of .555 only is not very high. This is hardly surprising since "all products" are not only composed of the few groups represented by the independent variables. The most significant predictor appears to be "clothing."

If the model of satisfaction composition is, in fact, correct, and respondents report their importance weighting of attributes truthfully, no significant deviations in the rank order of the mean importance weights and the standardized regression coefficients should be observable.

6. Differences in Consumer Satisfaction Among Occupational and Racial Groups.

Are there significant differences in consumer satisfaction between different socio-economic groups? To answer this question an analysis of variance was performed on MAC III scores for the effects of occupational status (A) as expressed through the differences among white-collar (A1) and blue collar (A2) workers; and for race (B), distinguishing between whites (B1) and blacks (B2).

The following patterns emerged:

1. Whites (B1) are significantly more dissatisfied with market goods than blacks (significance level of $\alpha = .01$)
2. White collar workers are significantly more dissatisfied with market goods than blue collar workers ($\alpha = .05$)

Turning to specific products, we note that:

3. Whites are significantly more dissatisfied than blacks with clothing ($\alpha = .001$)
4. Whites are significantly more dissatisfied than blacks with breakfast cereal ($\alpha = .05$)

These results are quite striking, particularly since they appear to be counterintuitive at first sight: Evidently, the higher socioeconomic strata are more dissatisfied because they are also more aware of the range of issues and possibilities associated with the market. Literary and general exposure to mass media may generate a predisposition to dissatisfaction. A more general expression of this phenomenon was observed by Strumpel for the same population: A different level of attainment in terms of a hierarchy of goals or needs seems to be evidenced by responses of different professional and racial groups, suggesting that blacks and blue-collar workers tend to be more concerned -- and more satisfied -- in the attainment of material goods.⁶ These results, it should not be forgotten, were obtained from a specific subpopulation - young families with employed family heads. Very different patterns may, in fact, be observed among the ppor, the aged, unrelated individuals, or the unemployed.

Product	White Collar/ Blue Collar/		White Collar/		Blue Collar/		Grand		F		
	White		Black		Black		Mean				
	Mean	Variance	Mean	Variance	Mean	Variance	A	B			
luncheon Meat	-3.62	1498.52	1.19	1301.14	5.78	1203.04	12.92	1111.33	4.07	0.80	2.51
breakfast Cereal	-5.03	1788.24	-1.13	1663.44	4.76	1108.32	18.95	1365.39	4.39	2.24	6.09*
clothing	-2.62	1245.85	-4.03	1315.74	9.75	795.98	24.84	1108.78	6.98	1.41	12.85***
all	-4.80	1459.89	1.50	1551.23	-0.24	914.22	18.34	1300.01	3.70	3.98*	2.94
sum	-3.84	873.38	-1.08	899.10	5.84	435.75	19.30	715.35	5.06	2.91	9.97**

* significant at the 0.05 level

** significant at the 0.01 level

*** significant at the 0.001 level

Table 9: Summary Results of the Analysis of Variance of Satisfaction Ratings of Different Professional and Racial Groups. (Based on MAC III Scores)

(In judging these results it should be kept in mind that the sample of respondents consisted of young families whose head of family was employed. This seems of particular significance for the black respondents. A more representative sample of the entire black population would most likely lead to different results.)

The analysis of variance results add further validity to the use of MAC III scoring: While the raw scores did exhibit patterns of variance which went in the same direction as the MAC III scores, their lower power of discrimination resulted in not significant differences. (A test for homogeneity of variance revealed that the F-test is appropriate to these data.)

Correlation analyses between consumer satisfaction scores and general personality, self-efficiency/fate control, educational level, and other socio-economic characteristics were carried out. A detailed description goes beyond the scope of this paper. On the whole, the ordered external variables (i.e., variables measured on an ordinal scale) and the unordered external variables (i.e., variables expressed on a nominal scale) showed very weak patterns of association with mean consumer satisfaction scores. There may be, however, significant differences among subgroups defined on the basis of these variables. In any case, our results indicate that consumer satisfaction is relatively independent of attitudinal and general personality factors. This result strengthens the meaningfulness of these indices. We can now be more sure that we are not measuring general optimism or pessimism, or other general traits, when we pose questions on satisfaction with market goods.

V. Conclusion

1. A Single Aggregate Measure or a Profile of Consumer Satisfaction?

The discussion of aggregation and prediction showed that even within a product using scores that maximize homogeneity, no perfect prediction could be achieved. This would indicate that, while a one-figure measure is very compact, it entails by necessity a great deal of information loss. An aggregate figure remaining constant from one time period to the next may conceal opposite movements of components. This paper did not deal at all with a possible use of this index as an indicator of consumer riots or boycotts. Extreme dissatisfaction with a product or product group may result in rather drastic action on the part of many consumers. Greater familiarity with a measure of subjective well-being may encompass the knowledge of a "critical value" or a "red zone". These are ranges of very low satisfaction scores that are indicative of imminent trouble. An aggregative measure may balance such "critical values" or trouble indicators by improvements in other unrelated areas.

This shortcoming may be circumvented by the use of a profile of consumer satisfaction measures. A profile involves a set of measures rather than a single value. For each group of products, a separate index would be computed.

A profile may be of further use in the composition of separate aggregative indices for different socio-economic groups, since it is likely that distinctly different groups would not attach the same relative importance weights to different products or product groups. Different weights in the averaging process may well be indicators of changes that favor one group at the expense of another group.

2. Aspects of Longitudinal Studies of the Index of Consumer Satisfaction

Longitudinal studies measure changes in a phenomenon over time in absolute or relative terms. The Consumer Price Index measures percentage changes in the price level in terms of relative deviation from the level which was set equal to 100 in the base period.

This study investigated three different measurement scales of satisfaction. Each one would have different advantages and disadvantages for use in longitudinal studies.

Raw scores allow comparison in absolute levels of satisfaction between products. They are, however, not very sensitive to extreme values -- and thus to change.

An index using this scoring system (or any similar one, say, using scores 7, 6, 5, 4, 3, 2, and 1 instead of the reverse order applied in the present study) would lead to a single value measure of the form:

$$(4) \quad ICS_t = \frac{\sum_i W_i \bar{S}_{it}}{\sum_i W_i \bar{S}_{i0}} \times 100$$

where \bar{S}_{i0} and \bar{S}_{it} are the mean satisfaction scores across all individuals of the i^{th} product or product group at the base period and time t , respectively; w_i are the importance weights. Using the limited number of product groups investigated in this study and a more or less arbitrary set of weights for illustrative purposes:

<u>Product Group</u>	<u>Weight</u>
food	.30
housing	.38
appliances	.02
car	.10
clothing	.20

The denominator of this index would amount to: $(.3)(2.95)+(.38)(3.26)+(.02)(2.16)+(.1)(2.49)+(.2)(3.00)=3.016$.

Due to the fact that in these raw scores 1 stands for "very satisfied," an increase in satisfaction would be reflected in a smaller number. If, for example, satisfaction had fallen at time t for the same indicators to, say, a mean of 3.2, the index would be $ICS_t = \frac{3.2}{3.016} \times 100$

The actual number of categories and the size of the intervals between adjoining points may have to be chosen differently, so as to provide a sufficiently sensitive indicator. Until at least two different amounts have been computed, it will not be known how sensitive a certain scoring system is.

Using optimal weighted (MAC III) scores, or optimal monotone (CM III scores) results in a slightly different index. These two scoring systems will not reflect different absolute levels of satisfaction between variables since each item would have a different scoring system. However, they would discriminate better for extreme values, particularly if the variables have unimodal distributions. Since each variable has a zero mean, the weighted average satisfaction of the base period would be a weighted sum of zeros, which tantamounts to zero. Evidently a ratio of new to old mean satisfaction cannot be computed in this case. An index of the form:

$$(5) \quad ICS_t = \sum W_i \bar{S}_{it} - \sum W_i \bar{S}_{io} = \sum W_i \bar{S}_{it}$$

would have to be used. (The symbols have the same meaning as above, only \bar{S}_i are measured in terms of MAC III or CM III scores). In both time periods the same scores would be used. In other words, once a scoring system has been computed in the base period it is retained for subsequent periods.

The disadvantage of such a system is that it cannot be compared to the Consumer Price Index as readily as the 100-based system. A remedy can be found in changing all scoring systems by adding a constant to all scores, e.g., let $\bar{S}_i^* = \bar{S}_i + 100$, then the interval sizes are retained and an index of the following form results:

$$(6) \quad ICS_t = \frac{\sum W_i \bar{S}_{it}^*}{\sum W_i \bar{S}_{io}^*} \times 100 = \sum W_i \bar{S}_{it}^*$$

CM III scores established for a base period could be retained for several years till a large shift towards more outlying (negative or positive) mean satisfaction scores has occurred, suggesting a changing group perception in distances between adjoint points on a satisfaction scale. This change would be desirable to provide, in a possibly changing situation, again for a more sensitive instrument. If the index were to be used, amongst other applications as a crisis indicator this sensitivity to extreme values would be of utter importance.

The significance of changes in the index can be tested by F-or t-tests, as suggested by the results of the present study.

MAC III scores, for the reason suggested earlier, seem to be less desirable and useful in the construction of an Index of Consumer Satisfaction.

This study shows that in principle, an index of consumer satisfaction can be constructed, which will

- (1) reflect relative changes in satisfaction;
- (2) allow for an aggregation of satisfaction measures across
 - (a) attributes to products,

- (b) products to product groups, and
 - (c) product groups to overall satisfaction;
- (3) provide for sufficient sensitivity, to diagnose
- (a) differences in satisfaction between groups and thus, by implication
 - (b) between time periods.

The usefulness of such an instrument for the formulation of public and business policy and as market performance measure will, of course, increase with extended measurement across population and time, and with the experience in the use of the scoring systems discussed.

Footnotes

1. This paper reports the interim results of an ongoing study of consumer and citizen satisfaction. The research was financed by Contract No. 12-17-D5-1-6107 from the U.S. Department of Agriculture.

I would like to express my appreciation to Professor Martin Pfaff, who was instrumental in the conception and design of this project; to Professor James Lingo, without whose cooperation the development of scoring systems and aggregation procedures would have been impossible; to Dr.

Terry Cooper and Dr. Lily Huang, who performed the computer analysis; and to Mr. William K. Jackson who assisted in the design, administration and analysis of pretest questionnaires of the Indexes of Consumer and Citizen Satisfaction.

An earlier version of this paper was presented at the 84th Annual Meeting of the American Economic Association jointly with the Association for the Study of the Grants Economy, New Orleans, December 27, 1971.

2. Assistant Professor of Quantitative Methods, Wayne State University; Wissenschaftlicher Rat und Professor, University of Augsburg.
3. For detailed descriptions of the Consumer Price Index see The Consumer Price Index. A Short Description, 1967. U.S. Department of Labor, Bureau of Labor Statistics, and, Consumer Price Index: History & Techniques, Bulletin No. 1517, U.S. Department of Labor, Bureau of Labor Statistics.
4. James Lingo, "The Multivariate Analysis of Qualitative Data" Multivariate Behavioral Research, January, 1968, Vol. 3, No. 1, pp. 61-94.
5. η^2 measures the square of the average correlation of the scaled variable with all other variables in the system. It is used as a measure of homogeneity.
6. Burkhardt Strumpel, "Economic Life Styles, Values, and Subjective Welfare -- An Empirical Approach" paper presented at the joint meeting of the American Economic Association and the Association for the Study of the Grants Economy, New Orleans, December, 1971; in Eleanor Sheldon (ed.) Understanding Economic Behavior, New York: Lippincott, 1973.

References

- [1] Bradburn, N. M. & Caplovitz, O. Reports on Happiness, Chicago, Ill.: Aldine, 1965.
- [2] Bradburn, N. M. The Structure of Psychological Well-Being, Chicago, Ill. Aldine, 1969.
- [3] Guttman, L. A General Nonmetric Technique for Finding the Smallest Coordinate Space for a Configuration of Points. Psychometrika, 1968, 33, pp. 469-506.
- [4] Green, P. E. & Carmone, F. J. Multidimensional Scaling and Related Techniques in Marketing Analysis. Boston, Allyn and Bacon, Inc., 1970, XV + 203.
- [5] Lingoies, J. C. An IBM 7090 Program for Guttman-Lingoies Smallest Space Analysis-III. Behavioral Science, 1966, 11, pp. 75-76.
- [6] Lingoies, J. C. An IBM 360/67 Program for Guttman-Lingoies Conjoint Measurement-III. Behavioral Science, 1968, 13, pp. 421-b2.
- [7] Lingoies, J. C. A General Nonparametric Model for Representing Objects and Attributes in a Joint Metric Space. In: Jardin, J-C., (ed.) Archeologie et Calculateurs, C.N.R.S., Paris, 1970, pp. 277-98.
- [8] Lingoies, J. C. Some Boundry Conditions for a Monotone Analysis for Symmetric Matrices. Psychometrika, 1971, 36, pp. 195-203.
- [9] Lingoies, J. C. The Guttman-Lingoies Nonmetric Program Series, 1972, (in press).
- [10] Lingoies, J. C. A General Survey of the Guttman-Lingoies Nonmetric Program Series. In Shepard, R., Romney, A. K. & Nerlove, S., (eds.), Multidimensional Scaling: Theory and Applications in the Behavioral Sciences, Seminary Press, 1972, (in press).
- [11] Lingoies, J. C. & Cooper, T. PEP-I: A Fortran IV (G) Program for Guttman-Lingoies Nonmetric Probability Clustering. Behavioral Science, 1971 16, pp. 259=261.
- [12] Lingoies, J. C. & Guttman, L. Nonmetric Factor Analysis: A Rank Reducing Alternative to Linear Factor Analysis. Multivariate Behavioral Research, 1967, 2, pp. 485-505.
- [13] Lingoies, J. C. & Roskam, E. A Mathematical and Empirical Study of Two Multidimensional Scaling Algorithms. Michigan Mathematical Psychology Program, 1971, 1, pp. 1-169.
- [14] Pfaff, M. Theories of Market Systems: Implications for the Measurement of Market Performance. Philadelphia, Pa.: Marketing Science Institute Working Paper, April, 1968.

- [15] Pfaff, M. & Lingo, J. C. Measurement of Subjective Welfare and Satisfaction. Paper presented at the 84th Annual Meeting of the American Economic Association jointly with the Association for the study of the Grants Economy, New Orleans, December 27, 1971
- [16] Pfaff, A. B. & Pfaff, M. Methods of Welfare Economics in the Measurement of Market Performance. Philadelphia, Pa.: Marketing Science Institute Working Paper P-51-7, May, 1969.
- [17] Pfaff, M. & Pfaff, A. Toward an Index of Consumer Satisfaction: The Synthesis of Benefit-Cost with Nonmetric Scaling Concepts as Basis for the Measurement of Market Performance. Philadelphia, Pa.: Marketing Science Institute Working Paper P-51-8, May, 1969.
- [18] Rakam, E., & Lingo, J. C. MINISSA-I: A Fortran IV (G) Program for the Smallest Space Analysis of Square Symmetric Matrices. Behavioral Science, 1970, 15, pp. 204-205.
- [19] Smith, P. C., Kendall, K. M. & Helin, C. L. The Measurement of Satisfaction in Work and Retirement: A Strategy for the Study of Attitudes, Chicago, Ill.: Rand McNally and Company, 1969.

IMPLICATIONS OF THE INDEX OF CONSUMER SATISFACTION FOR PUBLIC
POLICY PERTAINING TO MARKET PERFORMANCE

Charles R. Handy
U.S. Department of Agriculture

I

This audience is no doubt well aware of the growing concern for the quality of life experienced by the American population. This concern is manifested in the rapid increase of interest in developing social indices or methods of accounting that will monitor the direction and rate of change of the many variables that can influence current and future national well-being. Interest in this activity has been expressed at all levels of government. Government agencies, consumer groups, educators, researchers, and many representatives of the business community have urged a larger effort to be expended in the development of social indicators, including measures reflecting changes in consumer satisfaction, aspirations, and expectations.

It seems to me that the social indicators under development can be broken down into two basic categories: (1) those measures or sets of data which monitor changes in the physical aspects of our environment, and (2) psychological or subjective measures which monitor how our society, or subsectors of our society, perceive their quality of life. The second category contains measures of attitude. They attempt to describe over time, how individuals and groups perceive their opportunities, satisfactions, and expectations. To date, the majority of manpower and financial resources has been devoted to developing the first category of social accounts.

The Department of Health, Education, and Welfare's document, Toward a Social Report, encouraged the development of these types of indicators and sketched the form they might take. These indicators would include: monitoring the level and trends of various pollutants and their effect on our natural resources; monitoring the quality and availability of education at all levels of society; more detailed monitoring of trends in criminal acts, and data on our penal institutions; monitoring trends in physical and mental health between different segments of our population; and monitoring trends in social mobility.

Social indices providing time series information on these issues, combined with our numerous and highly developed economic indicators, would greatly facilitate our ability to assess trends in our national well-being. Yet both of these types of social and economic measures omit a crucial variable in any assessment of well-being; the subjective attitudes of society. We need to "flesh out" both our economic indicators and social indicators with subjective or psychological indices reflecting a person's perception of a situation as it affects him. This brings us to the Index of Consumer Satisfaction and the second category of social indices.

When turning to subjective or psychological indicators, we move into an area that is highly experimental. There are very few examples of subjective economic indicators that are actually in use. This is understandable. These measures present tremendous measurement problems, problems of aggregating responses between individuals, groups, and products, and problems of interpretation. Another factor tending to restrict development of these indicators is that as researchers and economists, we have shown a strong affinity for precision of methodology and

concepts even if it means restricting areas of investigation. The rapidly changing social fabric of our society creates a vacuum of information concerning our values, attitudes, expectations, and satisfaction. Measures which document this type of information may provide a useful and timely empirical data base for anticipating and responding to real world problems.

The papers presented this morning focus on the conceptual and methodological problems of constructing indices of consumer satisfaction. This has been a necessary first step. Up to now the ICS was a promising but untested concept. Research reported this morning has shown the powerful techniques of nonmetric scaling can be successfully and meaningfully employed in the development of subjective measures of market performance.

Now that procedures for constructing indices of consumer satisfaction have been largely developed, I feel greater involvement from business firms, trade associations, government agencies, consumer groups, and other potential users is justified. As we move forward to operationalize the ICS, primary effort must shift from methodology questions to concern with maximizing useful information for public and private decision makers. Statistical and mathematical elegance is not enough for continued support. Thus, it is to the potential applications of the ICS that I would like to direct my remarks.

II

The Department of Agriculture became involved in developing an index of consumer satisfaction following a detailed evaluation of traditionally used measures of market performance such as value added, productivity, profit rates and market structure variables. This investigation concluded that while these measures were extremely useful, they all focused on the input side rather than the output side of the market. While we can measure the growth of our economy, efficiency of resource use, and product flows, we have virtually no information indicating the extent consumers feel the products and services available to them actually conform to their preferences and needs. Yet in our economy, consumer satisfaction is one of the ultimate goals of marketing activity. Considering today's environment of rapid change, various tabulations of the increased number and variety of products offer little evidence as to how well these market alternatives match a heterogeneous demand.

The demand for consumer satisfaction information from both the public and private sector is increasing. Knowing how the various segments of the population feel--their attitudes and frustrations--is increasingly necessary for the development of far-sighted corporate and public policy. Sources of consumer dissatisfaction are not static. Rather they require periodic monitoring since they may stem from changes in consumer values, attitudes, and expectations as well as from physical characteristics of products and services.

Both government and the business sector need to improve their dialogue with consumers. An index of consumer satisfaction could be a vital mechanism to facilitate this dialogue. Just the mere collection and publication of consumer satisfaction data may likely be beneficial since individuals tend to be more sensitive and responsive to those factors that are regularly measured. At present, this dialogue is limited largely to random complaints initiated by consumers, information obtained at hearings called by various government agencies, and private consumer research activity which is highly concentrated among a limited number of products and brands.

I now would like to mention several potential uses of consumer satisfaction information by the public and private sectors. A primary mission of the Economic Research Service is evaluating market performance and determining ways in which the marketing system can be improved to more effectively meet the needs of farmers, market intermediaries, and consumers. Consumer satisfaction and the responsiveness of markets to consumer needs are major dimensions of market performance that have received inadequate attention. Indices of consumer satisfaction with a market basket of food products and services would help close this gap in market intelligence.

Indices of consumer satisfaction would supplement traditional market signals from consumers to producers. This type of information could help commodity groups and marketing firms tailor their products and services to the expressed needs of different consumer groups. It would also help identify market voids currently unserved by available products or services. While this information would be available to all, it should be particularly useful to smaller firms and to agricultural commodity groups since they often do not have the resources or organization to conduct their own consumer research.

Another function of government is to be a resolver of conflicts. Government is concerned with promoting the welfare of its people and how to best plan and adjust policy to take care of citizen needs and wants. In carrying out this function, policy makers require a solid empirical information base for action. Public officials must be able to identify and anticipate citizen concerns and problems. Without regularly monitoring a large cross section of our population, it is very easy to be misled as to the actual state of consumer dissatisfaction.

Congress and local legislative bodies in recent times have frequently found themselves in a position of being called upon to act on consumer issues without the benefit of an adequate information base. The pressure for action may not allow enough time to plan and conduct appropriate research. Information from an ongoing index of consumer satisfaction would provide legislators a continuous objective data base to help guide decision making. This information may also provide evidence for evaluating the extent past legislation such as "Truth-in-Packaging" has been successful in reducing consumer discontent with this aspect of marketing.

Discussions with representatives of the Office of Consumer Affairs and the Federal Trade Commission indicated several areas where information from an index of consumer satisfaction could be useful. The Office of Consumer Affairs (OCA) receives thousands of letters per month containing consumer complaints. These complaints are categorized by topic, tabulated, and stored. The primary problem in relying on this information is that OCA has no way of knowing how representative it is. People who write complaints tend to fall into two groups, according to OCA. One group tends to have a lot of time on their hands, indicating a high proportion of retired persons. The other group tends to be very articulate, signifying above average education and income. An index of consumer satisfaction covering all product group classes would help OCA estimate how representative are various complaints received by them by showing how satisfactions or dissatisfactions with products and services are distributed across specific segments of the population.

Representatives from the Federal Trade Commission suggested that this type of index could be of use to them in: (1) the allocation of resources to important problem areas as perceived by consumers; and (2) the development of programs and regulations which reflect differences in the perceptions, needs, and predispositions of different segments of the population.

Selection of trade or marketing practices to challenge relies heavily on individual complaints, suggestions, and staff investigations. An index of consumer satisfaction again could indicate the extent individual complaints are representative of all consumer groups. Information from indices of consumer satisfaction could also aid the Commission in designing programs to take account of individual product and consumer differences. The proposed index is able to identify particular product attributes causing dissatisfaction. For example, if dissatisfaction for a particular product group was not related to price, the Commission may be misallocating its resources in mandating industry action to improve price information in these particular product lines.

Index of consumer satisfaction information could be of additional use to the private sector by the encouragement of voluntary action and self-regulation of business practices. Legislation and regulations once enacted, tend to impose universal behavior and values on all. Voluntary action enables business firms to retain greater flexibility in responding to the many minority interests and market segments they seek to serve.

The Council of Better Business Bureau recently has become more active in encouraging business self-regulation. Plans are to set up a national consumer information bank. Information from local bureaus and the national office will include data on: promotion practices, product performance, consumer attitudes, and complaints. An index of consumer satisfaction could serve as additional input and reference point for the data bank.

On a more general or macro basis, the ICS can be viewed as a social indicator reflecting over time changes in satisfaction with specific and general aspects of our market economy. At this stage, we can only speculate as to the stability of the indices over time. Increased experience with movements in the indices of product or service satisfaction may help identify or predict buying patterns, and may also be indicative of potential consumer protest should an index fall to a certain level of dissatisfaction.

These comments clearly indicate that in my opinion, subjective indices of economic performance have the potential for becoming important inputs affecting policy decision making. Subjective or psychological indicators provide an additional perspective or dimension to balance traditional economic, accounting, and engineering measures of performance.

Subjective indices are not a panacea. There will continue to be serious questions of interpretation as these measures come into use. Consumer satisfaction as a goal also has limitations. Individual consumer preferences in some cases may have to be restricted or disallowed if long run social costs outweigh private benefit. Generally, the technical feasibility of constructing subjective measures of economic performance has been demonstrated; they appear to provide worthwhile information; and they have a number of potential users.

THE INDEX OF CONSUMER SATISFACTION AND CORPORATE MARKETING POLICY

Robert W. Pratt, Jr.¹
General Electric Company

As I understand it, the Index of Consumer Satisfaction (ICS) is being developed to supplement "cost measures" provided by the Consumer Price Index with a "benefit measure"--that is, a "subjective" indicator of consumer satisfaction. To quote Dr. Pfaff, the ICS "tries to determine the benefit that consumers receive--or the degree of satisfaction they experience--from the operation of the market."

Consumer satisfaction has been a basic concern of the business community for a long time. In a real sense, it's what we sell. And I believe business will be warmly receptive to both methodological and substantive insights generated by this work. For at least the last twenty years, there has been a wide recognition among researchers and planners in industry, government and the academic that we must move beyond economic and demographic variables and develop the capability to identify and measure sociological and psychological variables relevant to a particular issue. Witness the recent proliferation of work in the general area of social indicators.

My assigned role this morning is to comment briefly on potential implications of the ICS for formulation of corporate marketing policy--not public policy, not broad business policy, but marketing policy.

While the ICS, in the form just described, may have immediate implications in the public policy area, I foresee difficulty in using it as an input to marketing policy. As I proceed to explain why, it will be useful if you picture yourselves as members of a marketing policy group within a single company. In general, you have available to you information about the consumer environment, the industry or industries in which you compete, competition, and, of course, your own products. The extent of this information varies greatly by product and industry. Your task is to make decisions that will optimize the market position of goods and services offered by your Company. Now, why do I feel that the ICS would have limited value in accomplishing this task?

First, consider the difference between the ICS and the Consumer Price Index. The distinction between an index structured using dollars as a common base and an index structured using measures of "satisfaction" as a base is an important one. In general, when making either tactical or strategic decisions, management will have available a great deal of extremely detailed information about costs and prices, and companies can and do take a wide variety of actions based on this information.

But I just don't believe that the Index of Consumer Satisfaction, in its present form, will offer comparable inputs for decision making. Although there may be a few exceptions to this, basic input data used to construct the ICS, regardless of how it is massaged and analyzed, is not sufficiently specific for use by any one company in the formulation of marketing policy. Let me use the appliance area as an example. The question asked on the ICS pilot study regarding satisfaction with household appliances was this:

"And what about the appliances you have, that is such things as your refrigerator, stove, and toaster, on the whole, how satisfied are you with your appliances?"

Answers were recorded on a seven point scale ranging from "very satisfied" to "not at all satisfied."

Now, wearing your marketing manager's hat, suppose the index is declining, indicating that consumers are growing less satisfied with their household appliances. This knowledge alone does not provide a basis for adjusting policies or programs. The management of a single company must have much greater detail before making decisions to change products or marketing programs. Perhaps I can clarify this point by describing one satisfaction/dissatisfaction study recently completed by General Electric. This will give you an idea of the level of detail required by management for decision purposes.

Over the past six years, we have conducted personal interviews with a national probability sample of 1,000 households each month. For four of those years, respondents have been asked the following question about each appliance available to them from among a list of thirty.

"Thinking about your (NAME OF PRODUCT), considering everything about it, which statement best describes your overall satisfaction or dissatisfaction with it?"

After considerable experimentation, we decided to use a five-point scale ranging from "extremely satisfied" to "extremely dissatisfied." For the first two years of this study, an open-ended "why" question was asked following every response. For reasons discussed below, the "why" question was later asked only for dissatisfactions. Over the four years, 48,000 respondents were questioned about 665,000 individual appliances. Here are some observations based on findings from this research.

1. The mean level of reported satisfaction across all appliances and all brands was 94 percent.
2. There was considerable variation in level of satisfaction across products and brands; also, level of satisfaction varied by such characteristics as age of appliance, time since most recent repair experience, satisfaction with that repair experience, and so on. In addition, as you might expect, the satisfaction level for a particular product or model frequently varies over time.
3. We found that respondents had a very difficult time articulating their reasons for satisfaction. Since, when acquiring a product, the buyer expects that he will be satisfied, the "why are you satisfied" question elicited such responses as: "It works," "It does what it's suppose to do," and so on. We eventually stopped asking reasons for satisfaction simply because the answers had no value for either operational or strategic decisions. Nothing was being learned that could be translated into marketing policy.
4. For an individual company to make decisions that will increase the benefits provided by its products to end consumers--and, hence, presumably increase their satisfaction--the company must have information specific to its products and models. Similar information about competitive offerings as well as a national Index of Consumer Satisfaction would, of course, represent potentially valuable norms. To realize this potential, marketing research would have to insure comparability between its measures and external measures, such as the ICS--for example, by using identical questions.
5. The information we have found useful for decision purposes is that associated with reasons for dissatisfaction, rather than reasons for satisfaction. For example, statements associated with breakdowns or repairs accounted for a high percentage of reasons given for dissatisfaction with all brands of television sets just a few years ago. Breakdowns and/or repairs accounted for a much lower percentage of all dissatisfactions for major appliances and

housewares. By far the dominant reason given for dissatisfaction with TV repairs was the fact that the repairman was not able to complete the job on one visit, whether because he was not qualified, did not have the right parts, or whatever. (Contrary to what you might think, for all appliances, cost of repair accounted for only 14 percent of the reasons for dissatisfaction.) Within GE, actions were initiated to improve both training of servicemen and availability of parts on repair trucks. One result is that the overall level of dissatisfaction with our repair performance has declined and the mix of reasons for dissatisfaction is changing. Incidentally, reasons for dissatisfaction are quite different for new as compared to older appliances.

6. In a number of instances, expressed dissatisfaction had nothing to do with the appliance asked about; rather, reasons referred to features and/or performance characteristics that the respondent would like to have. For example: "I'm dissatisfied with my refrigerator because it is not 'frost free'." This type of answer apparently reflects a form of "psychological obsolescence" with the present unit.
7. We learned that answer patterns are extremely sensitive to question wording. For example, if respondents are asked if there is anything at all disliked about an appliance, the recorded level of dissatisfaction jumps. Presumably one could design questions that would virtually eliminate expressed dissatisfaction. Our questions, of course, were designed to provide a level of detail needed for in-Company decision purposes.
8. Finally, in view of Martin Pfaff's earlier reference to studies that indicate a low correlation between job satisfaction and job performance, another finding that may be of interest to you is this: Respondents who said that they were dissatisfied with a specific GE appliance and also that they had no intention of replacing that appliance at the time of the interview, and who were found on a reinterview to have replaced the appliance, were just as likely to have purchased GE as similar respondents who initially said they were satisfied with the brand. While I believe we have explained this to our own satisfaction, the explanation is too lengthy to review here.

The essential point I have tried to make is that inputs to marketing decisions, if the results of these decisions are to prove truly beneficial to end consumers, must be relatively specific. This information will not be forthcoming from the ICS as presently constituted, if my understanding of the index is correct.

Let me make a few additional observations regarding the ICS that might serve as a basis for discussion.

1. First, only 342 respondents of the 574 sampled respondents completed the full set of questions, a loss of 40 percent. It would seem that any set of questions developed for use with a national sample, for purposes of developing a representative national index, would have to be simplified so that they could be answered by a larger percentage of designated respondents.
2. Second, the effort so far has paralleled as closely as possible the Consumer Price Index in both concept and execution. Given the difference in the nature of what is being measured, I would ask whether or not this is necessary or desirable?
3. Third, for some products, respondents were asked about the importance of various product attributes in their purchase decision. A great deal of research has been done that suggests that interpretation of answers to questions of this type must be made with great care. For understanding

purchase decisions, at least for consumer durables, I have become increasingly convinced that we come closest to actionable results when we have available panel data that can be analyzed using longitudinal techniques.

To conclude. In these brief remarks, I have attempted to focus on the potential for application of the Consumer Satisfaction Index to the formulation of marketing policy. Extensive reference to one GE study has been made simply to convey the level of detail management wants as an input to policy formulation. I have intentionally avoided reference to what I think are interesting and provocative issues associated with the ICS, such as the question of the extent to which government should be held responsible, or assume responsibility, for the "happiness" of citizens. While interesting, this type of question didn't seem to fit my assignment here. I have also avoided talking about specific methodological detail. Such discussion could go on all day and probably much longer, and I see no particular purpose to be served by making scattered observation.

I believe results from empirical research of the type described this morning will eventually become necessary inputs to policy making in both the public and the private sectors. The present effort is a well conceived and scholarly piece of research, and I hope the project will be continued.

As I stated earlier, measures of satisfaction/dissatisfaction have been of paramount importance to marketers for a long time. If they have not already done so, in the process of refining the design of the Index of Consumer Satisfaction, I would urge that its architects consider availing themselves of some of the knowledge that has been gained in this area over a considerable period of years.

Footnote

1. Manager, Business Research & Forecasting Operation, Major Appliance Business Group, General Electric Company.

THE ZEIGARNIK EFFECT IN ADVERTISING

James T. Heimbach

Nationwide Research Center, and
Jacob Jacoby, Purdue University

If advertising is to be effective, its impact should persist over time. It is not enough that an advertisement have some effect at the moment of presentation; rather, it must continue to exert its effect throughout the decision-making process.

In general, there are two broad approaches to enhancing the persistence of an advertisement's effect. First, the strength of the initial impact may be raised, thus raising the residual impact at any point along the decay curve. Second, the slope of the decay curve (i.e., the rate of loss of effectiveness) may be decreased. Either of these processes will lead to an advertisement presented at time t retaining greater impact at time $t + x$, the time of decision-making or purchase. The primary purpose of this investigation was to examine a technique for increasing the initial impact of an advertisement and the slope of the subsequent decay curve.

Numerous operational definitions of "advertising impact" have been developed. In general, they fall into three categories:

1. Memorial effect (e.g., awareness, recognition, recall, knowledge of product)
2. Attitudinal effect (e.g., opinion, liking, belief, preference, intention, conviction)
3. Behavioral effect (e.g., inquiry, purchase).

The pair of studies to be discussed in this paper focus on two types of memorial effects--awareness and recall. The assumption which underlies this investigation is: "That version of an advertisement which produces the greater memorial impact is the more successful version of the advertisement." Immediate measures (e.g., awareness, immediate recall) reveal changes in the initial impact of the advertisement, while delayed measures (e.g., awareness change, delayed recall) reveal the compound effect of changes in the initial impact and changes in the rate of decay of effectiveness. Both types of measures were employed in this investigation.

Since Zeigarnik's (1927) classic study, it has repeatedly been demonstrated that incomplete tasks are better remembered than complete tasks (cf. Butterfield, 1964). Zeigarnik's explanation for this effect (now named for her) was that the subject beginning a task develops a need to complete it. If he is prevented from doing so, he is left in a state of tension, which manifests itself in improved memory for the uncompleted task.

Extending this explanation from tasks to messages (i.e., advertisements), one could suggest that hearing the beginning of a message leads to the development of a need to hear the rest of it--rather like waiting for the second shoe to drop. The resulting tension leads to improvement in memory for that part of the message which has already been heard.

As might be supposed, this is a rather short-term phenomenon--one does not stay up all night waiting for the second shoe to drop. So what is its significance for advertisers? The key lies in the fact that, in the situation discussed, the subject is unable to complete the task on his own--he does not know the rest of the message. This is not the case with an advertisement. Few people hearing "Winston tastes good like a _____" would be unable to complete the message. Most probably, a good proportion of readers have completed it just now. This illustrates two further forces leading to improved recall for the message. The first, suggested by Heller (1956), is that completing the message is positively reinforcing, thereby improving learning and memory. A second force is provided by the fact that the audience will have actively participated in the message, and it is a well-known learning principle that active participation, as opposed to passive reception, improves learning and memory. These effects would be expected to persist over extended periods of time.

In addition to testing the general hypothesis (i.e., that incomplete ads would be better remembered than complete ads), the two experiments also examined other factors which could be hypothesized to strengthen or diminish the magnitude of the Zeigarnik effect.

1. Subject's need for achievement (nAch). Several studies (e.g., Atkinson, 1953; Weiner, 1966) have found that high n-Achievers show a stronger Zeigarnik effect than low n-Achievers.
2. Subject's prior familiarity with the advertisement. It is expected that greater familiarity will lead to better ability to complete an interrupted message, thus strengthening the Zeigarnik effect.
3. Subject's prior involvement with the product and brand advertised. It is predicted that subjects more involved and interested in the subject of the commercial will pay closer attention to it and will make more effort to complete an interrupted message.
4. Presence or absence of a jingle in the advertisement. The presence of a jingle should strengthen the Zeigarnik effect in three ways: (a) by adding strain for completion of the musical phrase to that strain already present for the verbal material; (b) by increasing the awareness of the interruption; and (c) by providing a well-known slogan which should be highly familiar and therefore easy to complete when interrupted.
5. Presence or absence of a post-commercial pause. Providing a brief pause between the termination of the commercial and the resumption of the program should enhance the Zeigarnik effect by giving subjects a chance to complete the commercial without interference.

Experiment I

Procedure

Two video tapes were prepared of a 30-minute television program, each containing four test and five filler commercials. These commercials were inserted into existing program breaks (i.e., replaced the original commercials appearing in the program as broadcast). Two test commercials (for cigarettes and soft drinks) were for highly familiar products, were highly familiar themselves (as established in a separate pre-test with different subjects),

and contained jingles. The other two test commercials (for rug shampoo and automotive services) were relatively unfamiliar and did not contain jingles.

One commercial of each type was presented in complete form in Tape 1 and incomplete--the last 5 to 6 seconds deleted--in Tape 2. In a cross-over control design, the other two were presented complete in Tape 2 and incomplete in Tape 1. This design allows analysis utilizing a replication factor to determine the reproducibility of the results.

Subjects were selected by administering the nAch scale from the Personality Research Form (Jackson, 1967) to 172 Introductory Psychology students at Purdue University during the Spring of 1970. The 30 highest and 30 lowest scoring students were selected for this experiment.

The two video tapes were shown simultaneously on 24-inch monitors in different rooms, with 30 subjects--15 high and 15 low in nAch--assigned to each room. Immediately following presentation of the program, the subjects completed an advertisement-recall measure. This measure consisted of asking the subjects to identify the product class (e.g., toothpaste) and the brand name and to provide a detailed description of the contents of each of the nine commercials. Upon completion, this measure was collected and an evaluation measure was distributed. This consisted of a listing of all the commercials, each accompanied by a seven-point, good-bad semantic-differential scale.

The recall measure was scored on a five-point scale. A score of zero represented total non-recall for a commercial, one point was assigned for recall of the product class, one point for the brand name, and zero, one, or two points were given for recall of specific copy points. The questionnaires were scored by two independent raters, who achieved a conspect reliability of .96.

The hypotheses tested in Experiment I were:

1. Incomplete versions of commercials will produce greater recall than complete versions.
2. High nAch subjects will manifest a stronger Zeigarnik effect than low nAch subjects.
3. Familiar commercials with jingles will produce a stronger Zeigarnik effect than unfamiliar commercials without jingles.

Results

Profiles of the recall scores are shown in Figure 1. It can be seen that the high-nAch subjects exhibited a Zeigarnik effect for both high-familiarity commercials with jingles (soft drinks and cigarettes) while the low nAch subjects showed the effect only for the cigarette commercial. The unfamiliar commercials without jingles consistently produced a small anti-Zeigarnik effect, in that they tended to be recalled less well when interrupted. The analysis of variance, however, revealed that neither the main effect for completeness nor the interaction between completeness and nAch were significant. The predicted interaction between completeness and "familiarity/jingle" was marginally significant at $p < .10$.

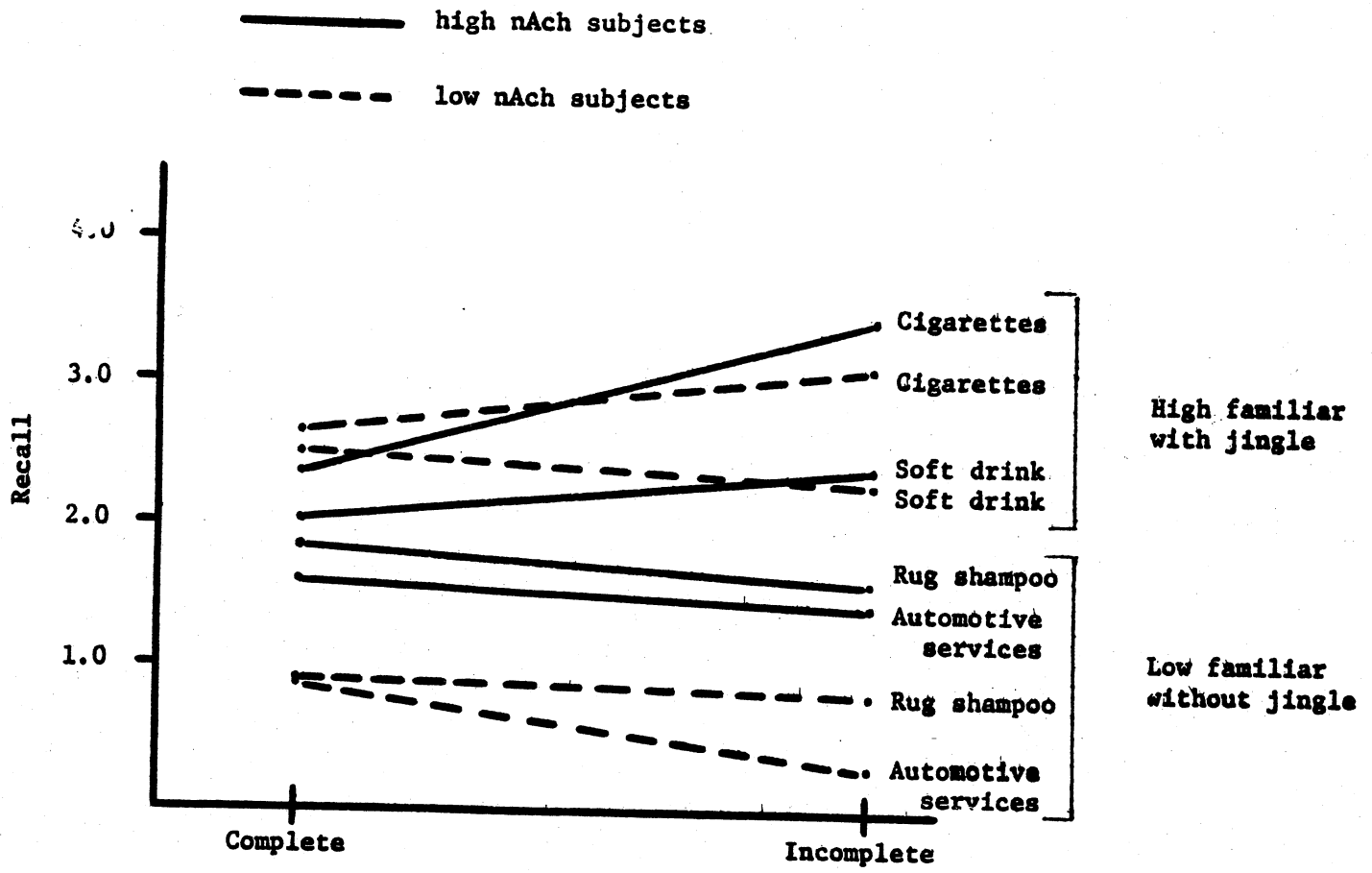


Fig. 1. Profiles of recall scores, Experiment I.

While it was not expected that the technique of interrupting a commercial would have any effect upon subjects' liking for it, this measure was included because it was felt that the possibility of either a positive or adverse effect should be explored. One significant effect of interruption upon evaluation did appear: a marginally significant ($p < .10$) triple interaction between completeness, nAch, and familiarity/jingle. This interaction may be due to an unplanned confounding factor. The two high-familiarity, jingle-containing commercials were for pleasure-oriented products (cigarettes and soft drinks), while both low-familiarity, non-jingle commercials were for functional products (rug shampoo and automotive services). It appears, then, that the high-nAch subjects preferred the complete versions of functional-product commercials and the interrupted versions of commercials for pleasure-oriented products, while the low-nAch subjects exhibited the opposite pattern, preferring complete pleasure-oriented and incomplete functional product commercials.

Although this experiment did not provide strong support for the application of the Zeigarnik effect to broadcast advertising, it was felt that since all results had been in the expected directions, the question did merit further study. Therefore, a second, somewhat more elaborate and better controlled experiment was designed to clarify some of the questions raised, as well as to explore new issues.

Specifically, the following questions were raised:

1. What is the effect of different levels of subjects' involvement with the brand and product advertised? Will those who use the product or brand pay closer attention to the commercial or make more effort to complete an interrupted message than those who do not use the product, thus showing a stronger Zeigarnik effect?
2. What is the effect of different levels of prior awareness of the brand advertised? Will subjects more highly aware of the brand pay closer attention to the commercial, thereby exhibiting a stronger Zeigarnik effect?
3. Will allowing a brief pause after the commercial strengthen the Zeigarnik effect by allowing the subjects a chance to complete the interrupted commercials without distraction?
4. Does the presence of a jingle enhance the Zeigarnik effect in the absence of large differences in the familiarity of the commercials?
5. What would be the effect of truncating the beginning of a commercial rather than deleting the end?
6. Can a Zeigarnik effect be shown for other memorial criteria than unaided recall of copy points? Specifically, can basic awareness of the brand name be stimulated via the Zeigarnik effect?
7. Finally--and most importantly--does the Zeigarnik effect persist over time?

The second experiment was performed to answer these questions.

Experiment II

Procedure

Six audio tapes were prepared of the sound portion of a TV show. Again, each contained four test commercials, two with and two without a jingle. In this experiment, however, all test commercials were of roughly equal familiarity. A pre test, in which 24 subjects viewed the four commercials (counter-balanced for order effects via a Latin square design) and assigned familiarity ratings from 1 to 7, established that the range of familiarity among the commercials was from 5.38 to 5.70; no differences were significant at the $p < .25$ level. The products advertised were cigarettes, chewing gum, mouthwash, and a headache remedy.

Each commercial was presented in one of six modes on each of the six tapes: it could be presented complete, interrupted at the end, or truncated at the beginning; in addition, it could be followed by immediate resumption of the program or by a four-second delay preceding program resumption.

One week prior to the presentation of the program, students in three classes of Introductory Psychology at Purdue University (Spring 1971 semester) were tested on their usage patterns and brand awareness for a number of low-price consumer products. This testing was performed by a graduate student having no ostensible connection with the present study. The awareness measure consisted of asking the subjects to list five brands of each product class. The product/brand usage measure asked if the subject used each product, whether he had a regular brand, and what that brand was. This pretest was used to select brands for use in the experiments as well as to determine which tape each subject would hear.

Each tape was played to a group of 15 subjects which included three subjects in each of the following five categories for each advertised brand: (1) regularly uses advertised brand; (2) uses product but not advertised brand and is highly aware of brand; (3) uses product but not advertised brand and is not highly aware of brand; (4) does not use product and is highly aware of brand; and (5) does not use product and is not highly aware of brand.

The subjects were told that the experiment was concerned with the relative amount of information carried by the audio and video portions of TV broadcasts. Immediate recall for the commercials was tested directly following the presentation, using the same recall measure as was used in Experiment I. The page containing this measure was the second of three pages of the questionnaire; the other pages asked questions about the program itself in order to continue the deception.

Two days later all subjects were retested on brand awareness; this testing was performed by the same graduate student who had administered the pretest under the pretext that he was checking its reliability. One week after the presentation, the advertisement recall test was again administered to measure delayed recall. At this time the deception was dropped and the subjects were debriefed.

The recall measures were scored by two raters, with an overall conspect reliability of .97. The awareness measure was scored in two ways. The first provided a measure of "change in awareness." This was achieved by subtracting

the awareness score on the pretest from awareness on the posttest, counting five for listing the advertised brand first, four for listing it second, and so on, down to zero for failing to list it at all. The second scoring method provided a measure of "first-brand awareness"--listing the advertised brand first on the posttest gave a score of "one," while anything else gave a score of "zero."

Thus, this experiment contained five independent variables: (1) Completeness (complete; cut-at-beginning; and cut-at-end); (2) Subject's Product/Brand Involvement (does not use product; uses product but not brand; uses brand); (3) Subject's Prior Awareness of Brand (high; low); (4) Jingle (absent; present); and (5) Post-Commercial Pause (none; four-seconds).

There were also four dependent variables: (1) Immediate Recall (measured within 10 minutes of the presentation); (2) Delayed Recall (measured one week after the presentation); (3) Change in Awareness (posttest measure two days after the presentation); (4) First-Brand Awareness (measured two days after the presentation).

It was anticipated that the dependent variables would react similarly to the manipulations, although first-brand awareness was expected to be the least sensitive measure.

The major hypotheses were:

1. Interrupting the end of a commercial is superior to cutting the beginning, which in turn is superior to presenting it in complete form.
2. Increased involvement with the product and brand advertised leads to a stronger Zeigarnik effect.
3. Increased brand awareness produces a stronger Zeigarnik effect.
4. The presence of a jingle strengthens the Zeigarnik effect.
5. The presence of a post-commercial pause strengthens the Zeigarnik effect.

Results

The primary hypothesis--that interrupting commercials produces greater memorial impact than presenting them complete--was strongly supported. This effect was significant beyond the $p < .01$ level for immediate recall, delayed recall, and change in awareness; it was in the expected direction, although not significant, for the first-brand awareness. In addition, the hypothesis that cutting the end of the commercial would be best, cutting the beginning second best, and leaving the commercial uncut worst was also strongly supported. This pattern was exhibited for all four dependent variables, and each interval was significant at the $p < .05$ level on the Tukey (a) test for immediate and delayed recall and for change in awareness. The profiles of these data appear in Figure 2.

As was noted earlier, it is important that the memorial impact of a commercial persist over time. Therefore, the relative performance of complete, cut-at-beginning, and cut-at-end commercials was examined over time (i.e., from

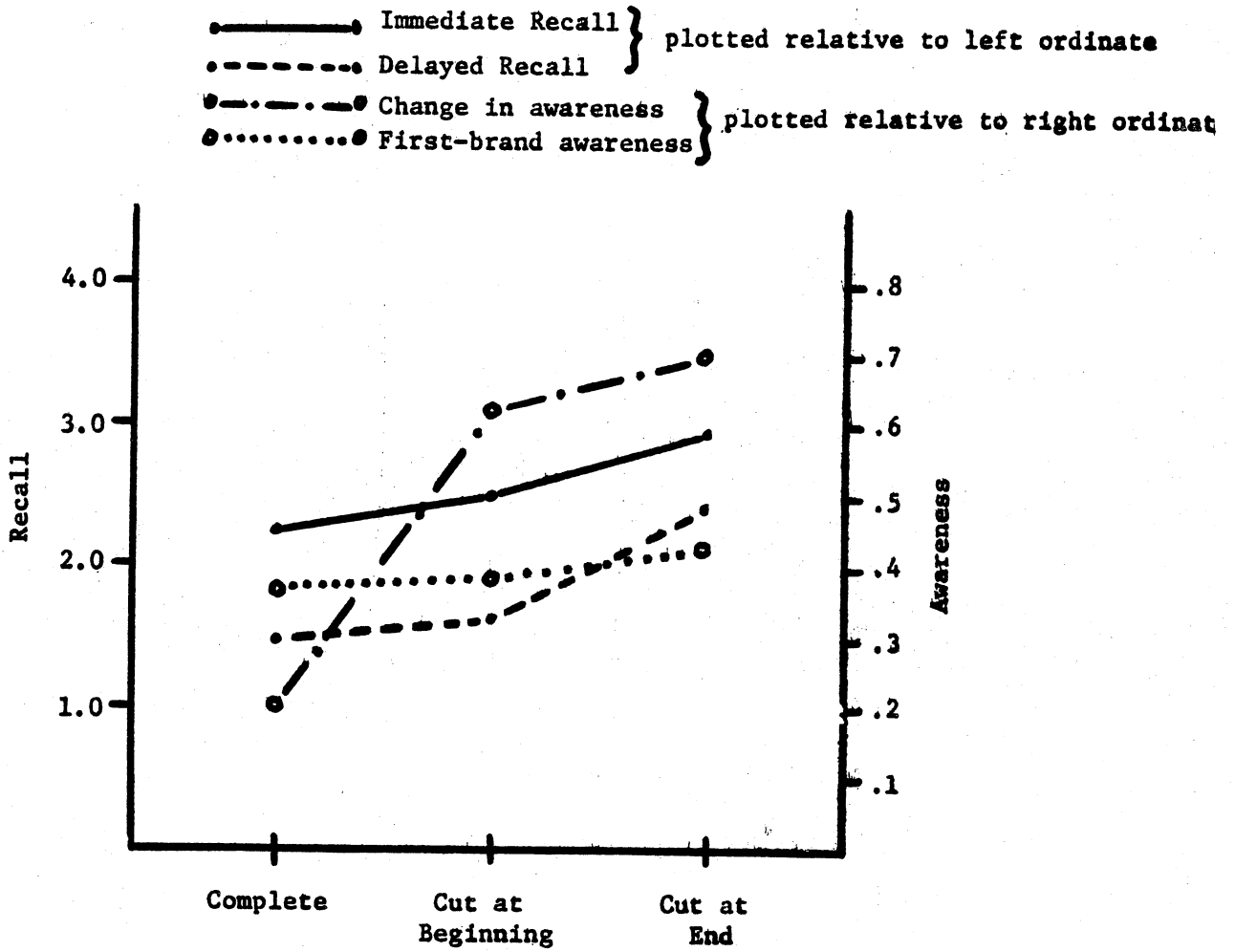


Fig. 2. Profiles of results for all dependent variables

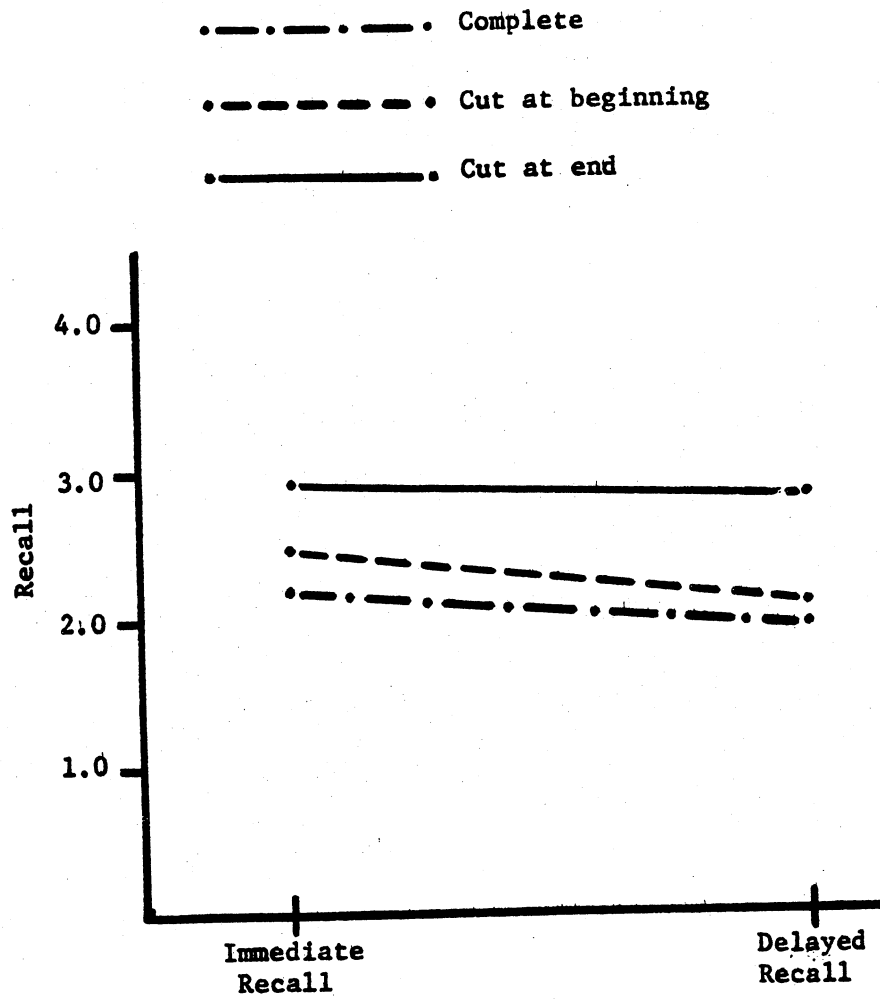


Fig. 3. Profiles of recall across time.

immediate to delayed recall--a period of one week). For the sake of simplicity, and since two points are not enough to establish the nature of the curve, a linear decay has been assumed. This plot appears in Figure 3.

It can readily be seen that cutting the beginning of a commercial produces an effect which decays over time, although it still produces recall superior to that for complete commercials after a period of a week. But the effect of cutting the end of the commercial appears to be a lasting one--its superiority over cut-at-beginning and complete versions in producing recall is even slightly greater after one week.

The hypotheses concerning the effects of product/brand involvement, brand awareness, and the presence of a jingle were not supported. None of these factors exerted any effect upon the Zeigarnik effect on any of the dependent variables.

The presence of a four-second pause following the commercial did have a significant effect, although this effect was considerably more complex than had been predicted. It appears that the presence of a post-commercial pause has different effects upon subjects with low and high initial brand awareness. As is shown by the profile of the triple interaction between completeness, post-commercial pause, and brand awareness (Figure 4), post commercial pause improved immediate recall of the cut-at-beginning commercials and decreased recall of complete commercials and those cut at the end for the subjects high in brand awareness. For those low in brand awareness, on the other hand, the post-commercial pause improved immediate recall of commercials cut at the end and decreased it for those presented complete or cut at the beginning. This same pattern also occurred at statistically significant ($p < .05$) levels on delayed recall and on first-brand awareness.

Discussion

It thus appears that a number of tentative conclusions can be drawn concerning the effects of the incomplete-commercial technique; however, a considerable number of questions remain.

First, it is apparent that the technique can have a substantial and significant effect upon recall for commercials, both immediately following the presentation of the commercial and one week later. In Experiment II, commercials cut at the end generated 33.8% more immediate recall than did complete versions of these commercials; one week later the recall advantage of the cut-at-end version was 52.4%. A similar effect was also observed for brand awareness.

It appears that truncating the beginning of a commercial is also an effective technique for increasing its memorial impact, although this technique is not as effective as deleting the end of the message.

While the data on familiarity of the commercial are not as complete as one would desire, it appears that the Zeigarnik effect will not occur with unfamiliar commercials.

There seems to be no support for the suggestion that the presence of a jingle will enhance the Zeigarnik effect, nor for the hypothesis that the viewer's involvement with the product and brand advertised will affect the strength of the effect.

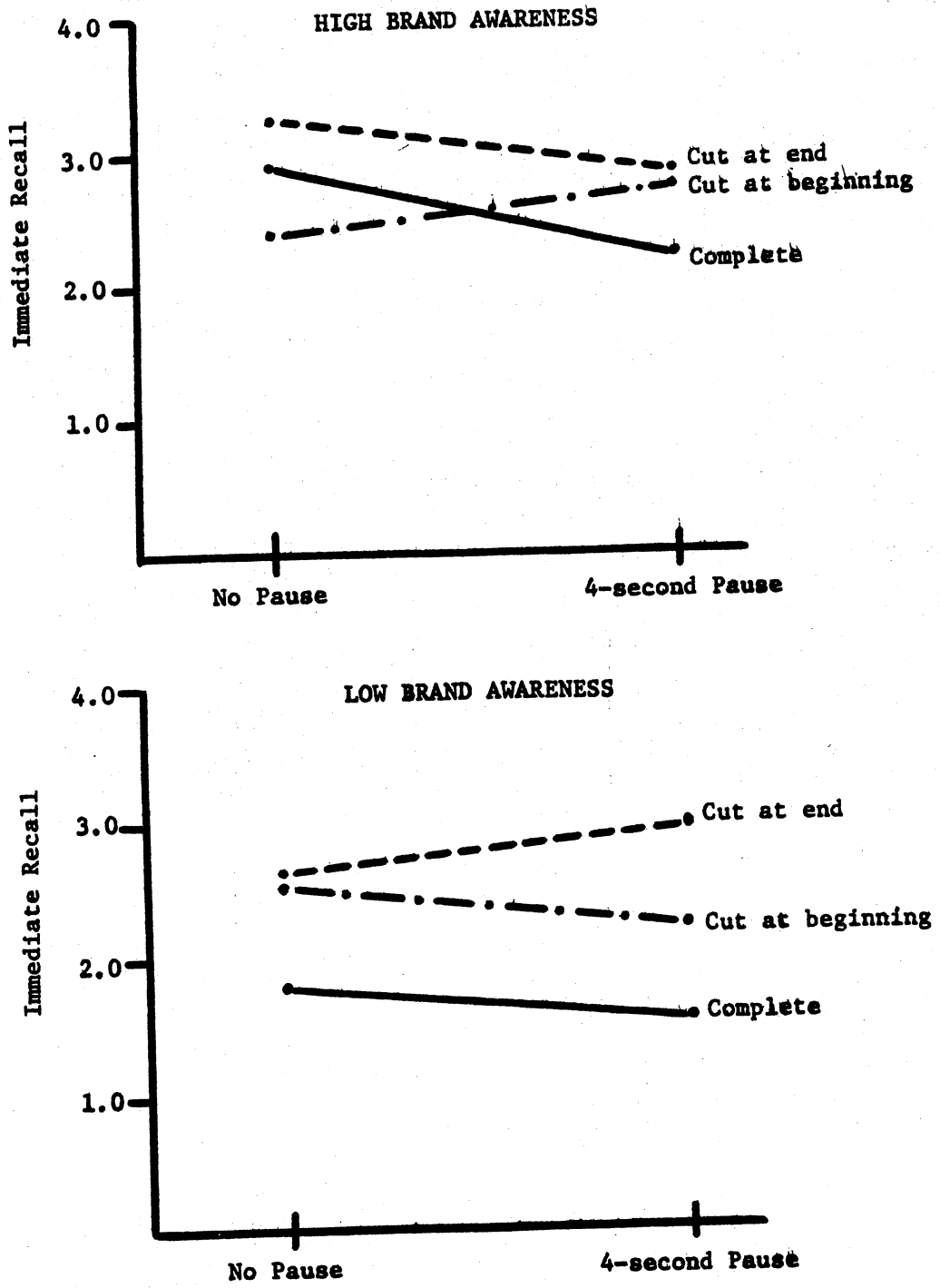


Fig. 4. Interaction between completeness, post-commercial pause, and brand awareness.

The question of providing a brief pause after the commercial obviously requires further research--the effects of this procedure are rather nebulous at this point.

Several basic questions remain. For example, to what degree is the improved recall for interrupted commercials due to a uniqueness or "shock" effect? Will a single interrupted commercial lose its effect with repetition, or would the production of a number of different interrupted commercials for different products cause the effect to disappear?

The effect of interruption of a commercial upon other than memorial criteria is also unknown. The single evaluation scale employed in Experiment I revealed the interaction which was interpreted--on a post hoc basis--in terms of functional vs. pleasure-oriented products, but did not reveal any main effect due to interruption. It must be emphasized, however, that this scale was probably not especially sensitive; the absence of significant effects suggests only that attitudinal effects are probably not very large. Further research on attitude, preference, and intention to buy as a function of interruption of commercials might well prove fruitful.

Finally, it should be noted that none of the advertisements utilized in these experiments were originally designed to be interrupted. It is probable that commercials could be specifically tailored to maximize the Zeigarnik effect; in addition, experiments could be performed with only the audio channel interrupted and the video channel left uncut. This technique also might prove effective.

In summary, it appears that the interrupted-commercial technique shows considerable promise for enhancing an advertisement's impact, but further research is necessary if the full potential of the technique is to be revealed.

References

- Atkinson, J. W. The achievement motive and recall of interrupted and completed tasks. Journal of Experimental Psychology, 1953, 46, 381-390.
- Butterfield, E. C. The interruption of tasks: methodological, factual, and theoretical issues. Psychological Bulletin, 1964, 62, 309-322.
- Heller, N. An application of psychological learning theory to advertising. Journal of Marketing, 1956, 20, 249-254.
- Jackson, D. N. Personality Research Form Manual. Goshen, New York: Research Psychologists Press, 1967.
- Weiner, B. Achievement motivation and task recall in competitive situations. Journal of Personality and Social Psychology, 1966, 3, 693-696.
- Zeigarnik, B. Uber das behalten von erledigten und unerledigten Handlungon. Psychologische Forschungen, 1927, 9, 1-85.

A STUDY OF MEDIA PREFERENCES
AND MEDIA CONSUMPTION¹

George H. Haines Jr.
David C. Efron
Graduate School of Management
University of Rochester

It is well known that there are different consumption rates of certain products (and brands) between demographic groups such as blacks and whites. There are, of course, several possible explanations for such phenomenon. However, investigations on the topic of race, such as Bauer, et. al (1965), Stafford, Cox, and Higginbottom, (1968), and Oladipupo (1970), have usually aimed at determining whether differential consumption patterns exist, not on why they exist. One study (Oladipupo, 1970) on Clairol hair coloring reported evidence that differential information sources were one of the causes of differential product consumption rates. This could imply, if generally true, that the market for consumer goods may not be efficient in an economic sense and that this lack of efficiency may cause consumption rate differentials.²

Further, studies that have attempted to explain consumption rate differences due to income and education (and some other demographic variables, but not race) on a basis of market segmentation have proved disappointing; for example, see Frank and Massy (1967). Again, one explanation for such results is that consumer goods markets are not efficient.

If this is the case, systematic differences across demographic groups in media consumption and program preferences within a media should be found.³ Therefore, this paper studies the proposition that there are systematic differences in media consumption and program preferences within media caused by race, income, and education.

Literature Review

Race, socio-economic class, and level of education have been found to have distinct effects upon individual consumption of and preference for media.

Samuelson, Carter, and Ruggels (1963) argued that education was a key variable in explaining media consumption. The argument presented was that as one's level of education increases, one's appetite for all media grows. At the same time participation in other activities grows. There being a fixed number of hours in a day, the end result is a person must give up some of the media consumption he would desire if there were no time constraint. Television loses the greatest amount of educated audience time. These people, SCR claim, substitute job-connected working hours, hobbies, organizations, and time with their children for some of the hours they would otherwise spend watching television. As education increases, media competing with television also gain. Better education brings better reading skills, and so magazines and books win readership at the expense of television's audience.

The authors present empirical evidence to support these assertions. Correlations between education and 1) job-connected hours per week, and 2) number of organizations active in both are significant at a Type I error of 0.05 and positive, while correlations between television and newspaper con-

sumption (in hours per week) and job-connected work hours per week are negative. The authors then adjust the data to remove the influences of jobs, volunteer organizations, hobbies, and dependent children. Partial correlation methods are used; the correlation between education and media now becomes positive for all media studied.

Media credibility has also been found to be related to educational achievement (Westley and Severin, 1964). The higher educated tend to trust newspapers more and the less educated tend to trust television more. "Social interaction," defined as participation in group activities, was also found to be related to media credibility by Westley and Severin (1964). Membership in three or more organizations is correlated (positively) with trust in newspapers, while the socially inactive put their faith in television. This finding corroborated the role involvement hypothesis of Samuelson, et. al. (1963).

Maxwell E. McCombs (1968), studying sources of political information, found race to be related to media consumption while education was important in explaining preferences for various mass communications material. Race, in part, explained the quantity of political information sought, but education was related to choice of media. Those of high school education or more, regardless of race, chose newspapers most often for political news; those who had not attained a high school degree were consistent viewers of television reports.

Gerson (1966) identified race as the dominant variable in media consumption patterns. This finding concurs with Carey's (1966) results. Carey argued the Negro experience was a set of factors which explained the significance of the race variable. Gerson argued that teenage blacks first join the dominant white sub-culture when they enter high school. Lacking preparation for the dating relationships common in this environment, they seek the mass media for advice. The media provide "...a mechanism...through which individuals learn to be motivationally and technically adequate in the performance of certain roles" (Gerson, 1966, p. 49). Therefore, one should find blacks using mass media, particularly television, in greater amounts than their white counterparts.

As can be seen, up to this point the research tradition had essentially studied demographic variables one at a time. Greenberg and Dervin (1970, 1970) attempted to rectify this weakness. The variable they added to race was income.

Greenberg and Dervin (1970, 1970) found that, although their sample interviewees differed when broken into low income (OEO designated areas; personal interviews used to collect data) and "general" populations (telephone survey used to collect data), the differences disappeared when the low income group was split black and white. For example, the low income sample spent 5.2 hours of the typical 16-hour working day watching television whereas the "general" population viewed for only two hours. Another difference lay in program preferences. Of twelve top-rated shows, a rank-order correlation between the low income sample and the general population sample is 0.03. However, when the low income sample was separated into racial groups no consistent differences appeared in program preferences. Similar findings are reported for media preferences for world news. The low income sample definitely prefer television, the general population is equally satisfied by newspapers and TV, and both black and white sub-samples of low income respondents prefer television.

Preference for local news produces a sharp contrast. The low income and general population samples named radio and newspapers as first choice, respectively. The low income blacks preferred people as sources of local news, but

whites of similar economic background were indifferent between TV, radio and newspapers. One possible explanation for this, advanced also by Honig, et. al. (1972), is a lack of news coverage in the inner city which creates dissatisfaction among blacks with the media's presentation of local news.

That the media, especially the broadcast media, do not program news with black interest in mind is further substantiated by Carey (1966). This paper, however, is in agreement with Greenberg and Dervin only on this point of news coverage. In fact, Carey (1966) generally contradicts all they say about the racial subsamples being not significantly different. He finds hardly any coincidence in preference among the top 40 television shows in 1963.

Carey (1966) concludes that the Negro experience forms a basis for black preferences which not only are different from white program and media preferences, but are also usually counter to the "cliche content" of most programs. Bauer and Cunningham reach the same conclusion: "Negroes are more concerned with matters close to their own life situation, and correspondingly they use the media proportionately more than do equal-income whites for recreation, diversion, and escape, and less as a way of maintaining contact with the realities of the world around them -- except for using advertising to learn what is going on in the market" (Bauer and Cunningham, 1970, p. 124). Bauer and Cunningham's figures show blacks to be heavier consumers of television and radio, lighter consumers of magazines and newspapers, than whites.

So the research findings end on a note of almost total disagreement between Greenberg and Dervin and the previous literature. There are three possible explanations for this:

- 1) Everyone else is wrong;
- 2) Greenberg and Dervin's results are idiosyncratic and cannot be replicated;
- 3) Greenberg and Dervin's results are replicable, but the conclusions Greenberg and Dervin draw from them are incorrect. This could be the case if (a) their analysis technique was a poor one for the problem at hand, and/or (b) their exclusion of the effects of education tended to mask racial effects.

Replication of Greenberg and Dervin

The Greenberg and Dervin Data

Two samples were taken by Greenberg and Dervin in their 1967 survey: (1) a "low income" sample consisting of random personal interviews in an area of Lansing, Michigan designated by the Office of Economic Opportunity (OEO) as having a high concentration of low income residents, and (2) a general population sample taken randomly from the Lansing telephone book. This latter sample was interviewed by telephone.

The Rochester Data

The data in the present study are drawn from two surveys performed in the Rochester, New York, area. The first, conducted during 1969, collected data from a random sample of residents of Monroe County, New York. The sampling plan was designed to collect a higher proportion of respondents living in areas where black people lived than from white ghetto areas. All sampling was done on a personal interview basis by a specially trained force of black interviewers.⁴

The 1971 survey was designed to be a random survey from Monroe and surrounding outlying counties. The sampling plan was designed to draw randomly and proportionately. Professional interviewers, all white, were used. All interviews were personally collected. However, there was an underrepresentation of blacks in the sample because the professional white interviewers tended to be unwilling to conduct in home interviews in certain inner-city areas. The effects of this selection bias will be discussed later. The present analysis uses only the data from Monroe County (the city of Rochester is entirely within Monroe County).⁵

Greenberg and Dervin, using the standard X^2 measure of association test, show that no significant differences exist between their low income and general populations on the number of working television sets, but do find significant X^2 values on color television ownership, hours television viewed yesterday, and the medium preferred for world news. However, when they split their low income sample into blacks and whites they found no significant differences at a 5% Type I error level. The replication analysis is concerned with these four questions.

The first question is whether the results are an artifact of the statistical test used. Therefore, Greenberg and Dervin's data was first re-analyzed treating the percentage of respondents in any category as an estimate of the parameter of a multinomial distribution. Differences across samples or race are then tested for by using a X^2 test of homogeneity of multinomial distributions (Potthoff and Whittinghill, 1966).⁶ Table I presents the results of this reanalysis of the Greenberg and Dervin data. The results with the homogeneity test statistical procedure are exactly the same, in a qualitative sense, as the results reported by Greenberg and Dervin (1970, 1970) using the traditional X^2 association test.

Tables II and III present replications of these tests with the 1969 and 1971 Rochester survey data. As with Greenberg and Dervin, respondents are divided into groups according to whether they live in the central city (Rochester) or in the surrounding suburbs. The samples of city residents are split according to race.

Something has gone wrong! The results certainly replicate Greenberg and Dervin's in the sense that low income blacks look like low income whites (or vice versa), but now there seem to be no significant differences between city and suburban residents, aside from a couple which could have occurred by chance. What could have happened?

One possibility might be the kind of breakdown used by Greenberg and Dervin. They split their sample geographically according to whether a person lived in a low income area or not. However, the Rochester surveys collected actual income data from respondents. One discovery which arose from this was that many people who live in low income areas do not have low incomes, and, obviously, there are some low income people living in the so-called medium or high income areas.

Therefore the analysis was redone, breaking on the income data directly rather than attempting to use place or residence as a proxy for income. The results, for the 1969 and 1971 data, are presented in Tables IV and V respectively. Everyone living in the city of Rochester and earning less than (or just) \$5,000 per year was compared to the suburban sample. The low income city sample was also split on race; all respondents who were neither black

Table I
Greenberg and Dervin [12] Data Reanalyzed
Using Multinomial χ^2 Homogeneity Test

	Low Income N = 281	General Population N = 206	White N = 150	Low Income Black N = 131
Number of working TV sets	0 8 1 177 2 96	6 120 80	5 99 46	3 79 49
	$\chi^2 = 0.769$; 3.54 d.f.		$\chi^2 = 1.175$; 3.55 d.f.	
Color TV	own 25 Don't own 256	45 161	14 136	12 119
	$\chi^2 = 15.05^*$; 1.86 d.f.		$\chi^2 = 1.901$; 1 d.f.	
Hours Viewed Yesterday	0 67 1/4-3 3/4 65 4-7 3/4 73 8 or more 76	84 87 25 10	35 37 43 35	33 26 31 41
	$\chi^2 = 60.64^*$; 5.62 d.f.		$\chi^2 = 3.07$; 5.95 d.f.	
Preference For World News	TV 194 Radio 45 Newspaper 42	78 58 70	102 20 28	89 28 14
	$\chi^2 = 43.93^*$; 3.75 d.f.		$\chi^2 = 5.48$; 3.93 d.f.	

* denotes significance at $\alpha = .05$

Table II
 Replication of Greenberg and Dervin
 1969 Rochester Survey Data

	City of Rochester Residents ("Low Income")	Monroe County Suburban Residents ("General Population")	City of Rochester Residents ("Low Income")
	White	Black	Black
No. of working TV Sets	0	1	0
	1	87	40
	2	97	54
	or more	(185)	(94)
		$\chi^2=0.275$; 1.587 d.f. +	$\chi^2=0.0$; 1.333 d.f. +
Own Color TV	Own	62	44
	Don't Own	123	50
		(185)	(94)
		$\chi^2=3.113$; 1.580 d.f.	$\chi^2=2.678$; 1.321 d.f.
Hours Viewed Yesterday	0	19	11
	1-3	9	25
	4-7	54	38
	8 or more	103	20
		(185)	(94)
		$\chi^2=30.06^*$; 4.62 d.f.	$\chi^2=2.54$; 3.69 d.f.

* denotes significance at $\alpha=.05$
 + denotes analysis excludes the case of zero working TV sets

Table III
 Replication of Greenberg and Dervin
 1971 Rochester Survey Data

	City of Rochester Residents ("Low Income")	Monroe County Suburban Residents ("General Population")	City of Rochester Residents ("Low Income")
			White Black
No. of working TV Sets	0 1 2 or more $\chi^2=4.29$; 2.03 d.f. +	1 30 53 (84)	1 13 43 (87) $\chi^2=0.017$, 1.08 d.f. +
Own Color TV	Own Don't Own $\chi^2=9.12^*$; 2.03 d.f.	54 30 (84)	2 9 (11) $\chi^2=0.332$; 1.08 d.f.
Average No. of hrs. today TV turned on	0 1-3 4-7 8 or more $\chi^2=0.735$; 3.97 d.f. +	1 10 38 35 (84)	0 0 4 6 (10) $\chi^2=0.253$; 2.10 d.f. +
Preference for News Source	TV Radio Newspaper $\chi^2=0.0$; 3.74 d.f. +	56 8 12 (76)	9 0 1 (10) $\chi^2=0.0738$; 2.01 d.f. +

* denotes significance at $\alpha=0.5$
 + denotes analysis excludes 0 categories

Table IV
 Replication of Greenberg and Dervin using
 Reported Income Data -- 1969 Survey Results

	Low Income (= \$5000) City of Rochester Residents	Monroe County Suburban Residents	Low Income (= \$5000) City of Rochester Residents
	White	Black	
No. of Working TV Sets	0	0	0
	1	40	7
	2 or more	54 (94)	4 (11)
	$\chi^2=2.43$; 1.82 d.f.†		$\chi^2=0.0$; 1.19 d.f.†
Own Color TV	Own	44	4
	Don't Own	50 (94)	7 (11)
		$\chi^2=8.64^*$; 1.83 d.f.	
Hours Viewed Yesterday	0	11	1
	1-3	25	2
	4-7	38	5
	8 or more	20 (94)	3 (11)
		$\chi^2=11.52$; 5.28 d.f. ($\chi^2 \leq .05$ with 5.28 d.f. 11.85)	
	$\chi^2=0.24$; 1.13 d.f.		$\chi^2=0.13$; 2.95 d.f.

* denotes significance at $\alpha=.05$

† denotes analysis excludes the case of Zero working TV sets

Table V
 Replication of Greenberg and Dervin
 using reported Income Data -- 1971 Survey Results

	Low Income (< \$5,000) City of Rochester Residents		Monroe County Suburban Residents	Low Income (< \$5,000) City of Rochester Residents	
				White	Black
No. of Working TV Sets	0	0	1	0	0
	1	18	30	12	6
	2 or more	8 (26)	53 (84)	6 (18)	2 (8)
		$\chi^2=3.50; 1.29 \text{ d.f.}^+$		$\chi^2=0.031; 1.65 \text{ d.f.}^+$	
Own Color TV Set	Own	8	54	7	1
	Don't Own	18 (26)	30 (84)	11 (18)	7 (8)
		$\chi^2=3.57; 1.28 \text{ d.f.}$		$\chi^2=1.123, 1.68 \text{ d.f.}$	
Hrs. Viewed Yesterday	0	0	1	0	0
	1-3	5	10	5	0
	4-7	11	38	8	3
	8 or more	10 (26)	35 (84)	5 (18)	5 (8)
		$\chi^2=0.0; 2.46 \text{ d.f.}^+$		$\chi^2=2.30; 3.18 \text{ d.f.}^+$	
Preference for News	TV	16	56	10	6
	Radio	5	8	5	0
	Newspaper	4 (25)	12 (76)	3 (18)	1 (7)
		$\chi^2=0.176; 2.46 \text{ d.f.}$		$\chi^2=.906; 2.67 \text{ d.f.}$	

* denotes significance at $\alpha = .05$

+ denotes analysis excludes the Zero categories

nor white are excluded on this racial comparison. These results show no significant differences except when color television ownership is compared across low income and Monroe County residents in the 1969 survey results. Everyone is alike! Greenberg and Dervin's claim that suburban people are different is not replicated in these results. Their claim that low income blacks and whites are alike is replicated, but this runs counter to all other findings in the research literature.

Before assessing the implications of these results one final question must be asked: are the samples the same? The low income samples on television ownership are, again taking a 0.05 level of Type I error (for whites, $X^2 = 8.41$, 3.49 d.f.; for blacks, $X^2 = 5.56$, 3.16 d.f.). The general populations differ on responses to this question ($X^2 = 12.19^*$, 3.49 d.f.). All samples differ on color television ownership, reflecting the increase in ownership of color television. Hours viewed yesterday are also all quite different; preferences for news are the same in the two low income samples but are different in the general populations.⁸

Now, what can be concluded? First, it appears that Greenberg and Dervin have really shown one gets different responses to questions on media use, consumption, and preferences if one uses telephone rather than personal surveys. Second, it appears that using homogeneity tests may not be a very good technique for studying the general problem. These tests are at once too general and too restrictive. Previous research suggests race, income, and education all effect media consumption and preferences. But homogeneity tests tend to push a researcher into simple comparisons; an example is the exclusion of education in the Greenberg and Dervin analysis. In this sense they are too restrictive. They are also too general: they ask "are there any differences," when past research has provided clear evidence for the direction differences should take if they exist. Simple homogeneity or association tests make no use of such important information.

The purpose of the next section is to investigate this issue further, using an appropriate analysis technique, one which will allow for simultaneous study of effects of race, income, and education, and for use of knowledge of the direction of the effect of these variables on media consumption and program preferences.

Analysis Method

The analysis model devised by James Coleman (1964) is used to see whether systematic differences in media consumption and media preference "caused" by demographic factors exist. Three independent demographic variables are examined: race, income, and education.

The basic structure of the Coleman analysis may be briefly outlined for the case of one independent variable (see Figure 1).⁹ There are two possible states a respondent may be in: that of being a light user of television, and that of being a heavy user of television. Whether a respondent is in one of these states depends upon the initial state of the respondent and, in Figure 1, the respondents race and random factors ϵ_1 and ϵ_2 . ϵ_1 is a random shock tending to lead a heavy user to being a light user of television; ϵ_2 is a random shock in the opposite direction. β_x is the effect being white has on creating a light user of television, α_x is the effect being black has on leading a respondent to be a heavy user of television. It is assumed $\alpha_x = \beta_x$; that is, the effects of race are equal, but operate in different directions. This is a useful and usual assumption in cross-sectional analysis; relaxing it in any meaningful way

Figure 1

Basic Structure of Coleman Analysis Model

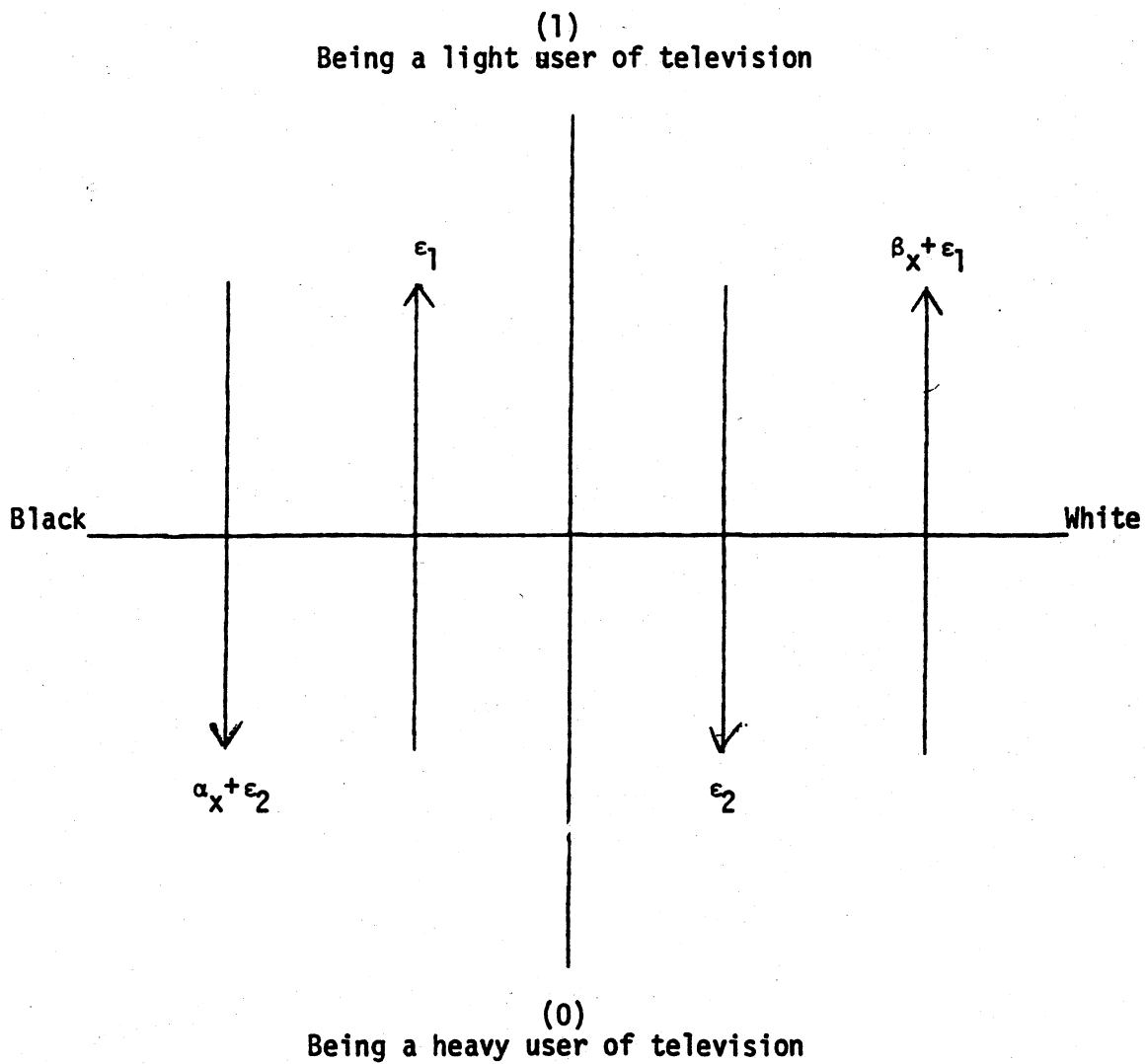


Exhibit I
Media Consumption Questions

1. Approximately how many hours a day is your TV set on?

Check Question ('69 survey only):

1a. About how many hours was your television turned on yesterday?

2. Are there times during the day when you would normally listen to the radio rather than watch TV?

3. Are there times during the day when you would normally read a newspaper rather than watch TV?

4. ('69 Survey only): Are there times during the day when you would normally go to the movies rather than watch TV?

5a. ('69 Survey only): Are there times during the day when you would normally read a book rather than watch TV?

5b. ('71 Survey only): Are there times during the day when you would normally read a book or a magazine rather than watch TV?

requires time series data.

Media Consumption

The questions asked and analyzed are presented in Exhibit I. A light user of television is defined as a respondent who answers less than ten hours to question 1; the same convention is used for question 1a. Since the analysis method yields an estimate of the direction of effects, alternative hypotheses were set down prior to running the analyses (in all cases the null hypothesis is that the demographic variables have no effect):

- (1) Higher educated respondents tend to be light users of television,
- (2) Higher income respondents tend to be light users of television, and
- (3) White respondents tend to be light users of television.

The proportion of respondents answering yes is taken as the dependent variable in the other analyses (instead of the proportion of light users). The null hypothesis remains the same; the alternative hypotheses are equivalent to the above (e.g., white respondents tend to answer yes to question 2).

A coefficient with a negative sign that is statistically unlikely indicates the demographic variable has an effect opposite to the alternative hypothesis specified above. An example of this is the effect of income on television consumption; that is, the results given in Tables VI and VII indicate that higher income respondents tend to be heavy users of television when effects of race and education are removed.

The results from the 1969 survey are presented in Tables VI and VII. In Table VI, A_1 , A_2 , and A_3 are the coefficients of Education, Income and Race respectively; U_1 , U_2 , and U_3 are standard normal deviates used to test the null hypothesis. R is a random shock toward the dependent state (for example, being a light user of television); S a random shock away from the dependent state. σ_{A_i} is the standard deviation of the coefficients in the analysis. Table VII is a "small sample" analysis. B_1 , B_2 , B_3 are the adjusted coefficients of Education, Income, and Race; U_1' , U_2' , U_3' are standard normal deviates used to test the null hypothesis. Tables VIII and IX are analogous, except these tables report results for the 1971 survey data.

The results given in Tables VI and VII indicate important racial effects. Race is significant in every category except "book rather than television." Income is also important. Education is never significant, but the sign is always in accord with the direction predicted by the theoretical argument of Samuelson, Carter, and Ruggels (1963).

There are some additional conclusions which these results would support. First, the coefficient values and significance levels in the two alternative forms of the television consumption question are quite close, indicating that either question is an equally satisfactory measure of television consumption. Second, it is instructive to compare the television consumption results in Tables VI and VII with those in Table IV. The results in Table IV show no significant differences between the hours viewed yesterday distributions. The difference demonstrates the ability of the Coleman analysis procedure to yield more information from a set of data than even the parametric frequency distribution homogeneity test employed in Table IV. Finally, the results in

Table VI
1969 Survey Results
Regular Analysis

	Education A1/U1	Income A2/U2	Race A3/U3	R	S	r_{Ai}
Hours a day TV set on N = 251	0.012 .204	-0.23 -3.97 ⁺	0.26 4.48 ⁺	0.67	0.29	0.059
Hours viewed Yesterday N = 250	0.038 .684	-0.26 -4.68	0.21 3.74 ⁺	0.74	0.27	0.056
Radio rather than TV N = 247	0.100 1.38	-0.011 -0.15	0.35 4.78 ⁺	0.31	0.25	0.072
Newspaper rather than TV N = 247	-0.050 -0.700	-0.020 -0.274	0.43 5.95 ⁺	0.44	0.20	0.072
Movie rather than TV N = 247	0.041 0.644	0.050 0.787	-0.23 -3.61 ⁺	0.30	0.84	0.064
Book rather than TV N = 247	-0.018 -0.258	0.23 3.26 ⁺	0.13 1.85	0.19	0.46	0.071

⁺denotes significance at $\alpha = .05$

Table VII
1969 Survey Results
Small Sample Analysis

	Education B1/U1'	Income B2/U2'	Race B3/U3
Hours a day	0.034	-0.257	0.652
TV set on	0.652	-5.42 ⁺	3.52 ⁺
Hours Viewed	0.063	-0.237	0.190
Yesterday	1.24	-5.32 ⁺	2.47 ⁺
Radio Rather	0.027	0.029	0.271
than TV	0.436	0.440	2.19 ⁺
Newspaper Rather	0.027	-0.039	0.437
than TV	0.488	-0.635	4.27 ⁺
Movie Rather	0.040	0.013	-0.22
than TV	0.773	0.230	-2.30 ⁺
Book Rather	0.067	0.24	0.078
than TV	1.05	3.70 ⁺	0.660

Table VIII
1971 Survey Results
Regular Analysis

	Education A1/U1	Income A2/U2	Race A3/U3	R	S	r_{Ai}
Hours a day TV set on N = 135	0.094 0.371	0.00069 0.00271	-0.024 -0.0955	0.88	0.05	0.25
Radio rather the TV N = 164	0.027 0.102	-0.18 -0.667	0.21 0.769	0.63	0.31	0.27
Newspaper rather than TV N = 155	0.23 0.868	0.078 0.290	-0.0020 -0.00733	0.55	0.14	0.27
Book or magazine rather than TV N = 168	0.31 1.16	0.15 0.540	-0.19 -0.724	0.45	0.29	0.27

Table IX
 1971 Survey Results
 Small Sample Analysis

	Education B1/U1'	Income B2/U2'	Race B3/U3'
Hrs. a day TV set on	--	--	--
Radio Rather than TV	0.12 ⁺ 1.91 ⁺	0.079 0.836	0.071 0.371
Newspaper Rather than TV	0.14 2.11	0.057 0.625	0.242 1.24
Book or Magazine Rather than TV	0.19 ⁺ 2.61 ⁺	0.19 ⁺ 1.96 ⁺	0.047 0.237

Table VII are very close to those in Table VI and are the same qualitatively. This indicates that the large sample analysis is adequate in this case.

The results from the 1971 survey are quite different. There are no significant coefficients in Table VIII; Table IX, the small sample analysis, indicates, an important education effect, with income being significant once. Race is never significant, but Table IX, at least has the expected sign (Sample size problems did not allow a small sample adjustment to be made in the case of television consumption.). The most striking feature of Table VIII is the size of the standard deviation of the coefficients compared to the size of the standard deviation of the coefficients in Table VI: roughly five times greater. The variances are significantly different. For example, comparing the variances for "hours a day TV set is on" yields an F value of 32.2 (134, 150 d.f.).

Earlier it had been noted that in the 1971 sample difficulties were encountered in getting the random sampling plan fulfilled because the professional interviewers were loath to work in inner-city areas. Now the effects of this can be seen. Many people doing commercial market research work seem to feel that low income urban areas can be excluded because, after all, "who needs data on some poor blacks anyway?" What these results show is the fallacy of such an argument, for what has happened is that the information content of the entire sample has been greatly reduced by the interviewing forces' actions.¹⁰

Television Program Preferences

The question asked and analyzed are presented in Exhibit II. The alternatives presented represent the actual programs scheduled on the three UHF television stations in Rochester in the spring of 1969; the nature of the alternatives leads to the supposition that there might be a strong racial effect on expressed preferences.

The analyses were run under the following alternative hypotheses (in all cases the null hypothesis is that education, income, and race have no effect):

- 1) as education level (A1) increases one tends to move toward the being interested in the program (health and birth control, for example),
- 2) as income (A2) increases one tends to move toward being interested in the alternative program, and
- 3) blacks (A3) are more interested than whites in the alternative program.

As before, a coefficient with a negative sign that is statistically unlikely indicates the demographic variable has an effect opposite to the alternative hypotheses specified above. R is a random shock toward preferring proposed program; S is a random shock away from preferring the proposed program. Γ_{A_i} is the standard deviation of the coefficients. These results are given in Table X. These results indicate a clear racial effect: the race variable is significant in every case except health and birth control. Education is never significant; in 3 of the 8 programs studied Income is significant.¹¹

Conclusions

Overall, what can be said? First that the time has clearly come to stop

Exhibit II
Television Program Preference Questions

1. On Sunday at 9 p.m. these shows are offered

_____ Bonanza
_____ Smothers Brothers
_____ Movie

If a program dealing with health and birth control information were presented, would you watch it instead?

(1) YES _____ (2) NO _____

2. On Saturday at 2 p.m. these shows are offered

_____ Wonderful World of Golf
_____ Happening
_____ College Sports

If a program dealing with Black History were also offered, would you watch it instead?

(1) YES _____ (2) NO _____

3. On Monday at 10:30 a.m. these shows are on

_____ Concentration
_____ Beverly Hillbillies
_____ Steve Allen

If, in addition, a local, inner-city amateur hour were offered, would you watch it instead?

(1) YES _____ (2) NO _____

4. On Monday at 10 p.m. these shows are on

_____ Carol Burnett
_____ Big Valley
_____ Movie

If, in addition, a program showing inner-city high school sports or Negro college sports were offered, would you watch it instead?

(1) YES _____ (2) NO _____

5. On Tuesday at 11:30 p.m. these shows are offered

_____ Johnny Carson
_____ Joey Bishop
_____ Movie

If a black drama program were also offered, would you watch it instead?

(1) YES _____ (2) NO _____

Exhibit II
Television Program Preference Questions (Cont.)

6. On Wednesday at 4 p.m. these shows are offered

_____ Movie
_____ Flintstones
_____ Dark Shadows

If an inner-city teen-age dance show were also offered, would you watch it instead?

(1) YES _____

(2) NO _____

7. On Thursday at 9 a.m. these shows are on

_____ Crossfire
_____ Captain Kangaroo
_____ Jack LaLanne

If a discussion program with well-known black persons were also offered, would you watch it instead?

(1) YES _____

(2) NO _____

8. On Thursday at 7 p.m. these shows are offered

_____ Truth or Consequences
_____ Merv Griffin
_____ News

If a Negro comedy program were also offered, would you watch it instead?

(1) YES _____

(2) NO _____

Table X
1969 Survey Results
Regular Analysis

	Education A1/U1	Income A2/U2	Race A3/U3	R	S	r_{Aj}
Health and Birth Control	-0.10	-0.021	0.23	0.507	0.39	0.17
Black History	-0.046	0.16	0.67	0.27	0.00	0.081
Inner City Amateur Hour	-0.567	1.99 ⁺	8.36 ⁺			
Black Sports	0.0058	-0.13	0.34	0.58	0.20	0.14
Black Drama	0.0406	-0.930	2.41 ⁺			
Black Team	-0.044	0.21	0.79	0.044	0.00	0.078
Dance	-0.565	2.701 ⁺	10.11 ⁺			
Black Discussion	-0.24	0.017	0.40	0.37	0.46	0.13
Negro Comedy	-1.82	0.130	3.02 ⁺			
	-0.015	0.18	0.76	0.0	0.083	0.083
	-0.180	2.14 ⁺	9.17 ⁺			
	0.015	0.20	0.45 ⁺	0.30	0.036	0.13
	0.114	1.54	3.44			
	-0.056	0.14	0.65	0.28	0.00	0.082
	-0.684	1.70	7.99 ⁺			

⁺ denotes significance at $\alpha = 0.05$. The sample size is 106 respondents in all analyses.

examining demographic variables' influence on consumer media use one at a time. The results plainly indicate that income, race, and education can all significantly effect media consumption. Similarly, income and race can both significantly effect television program preference. These are expected results given the initial "theory of market inefficiency" explanation of observed consumption rate differences.

The results also tend to reinforce the notion that Greenberg and Dervins' interpretations of their results are suspect. Finally, from a methodological viewpoint, the results show very clearly the importance of proper execution of random sampling plans.

Footnotes

1. The research reported in this paper was supported in part by the Systems Analysis Program, The University of Rochester, under Bureau of Naval Personnel Contract Numbers N00022-70-C-0076 and N00022-71-(-006) and in part by the Brookings Institution Workshop on Consumer Affairs at the University of Rochester. However, the conclusions, opinions, and other statements in this paper are those of the authors and not necessarily those of the Systems Analysis Program or the Brookings Institution. The authors are indebted to V. Srinivasan for helpful comments.
2. An efficient market is one in which market prices fully reflect all available information (Fama, 1970).
3. This paper labels the time a consumer spends watching television, etc., as consumption; hence the phrase "media consumption". However, watching television, etc., can be used to produce commodities, in the sense of Becker (1965), by combining inputs of goods and time; hence the view of time spent on media as being pure consumption is incorrect. Indeed, the Becker (1965) theory is a necessary link to explain why differential media use could be a reflection of market inefficiency.
4. The collection of these data was supported by a grant from the Consumer Research Institute, Inc., and would not have been possible without the aid of Marcus Alexis and Leonard S. Simon. We wish to express our gratitude to all the above, for without the data this paper would not have been possible.
5. The collection of these data would not have been possible without the aid of James Peck and Daniel Braunstein. Once again we wish to express our gratitude for this help.
6. Copies of a BASIC program for performing the computations of this test are available from the authors on request.
7. This is really not a very startling discovery. The reader should also see Harrison (1972, p. 26), where frequency distributions of male weekly earnings in March 1966 by race by residential area type are presented.
8. The results are as follows:

Color Television Ownership:	
General Population	$\chi^2 = 35.86^*$, 1.95d.f.
Low Income White	$\chi^2 = 90.37^*$, 1.99d.f.

- | | |
|-------------------------|-----------------------------|
| Low Income Black | $X^2 = 15.74^*$, 2.02d.f. |
| Hours Viewed Yesterday: | |
| General Population | $X^2 = 103.47^*$, 5.76d.f. |
| Low Income White | $X^2 = 25.09^*$, 5.91d.f. |
| Low Income Black | $X^2 = 34.28^*$, 6.02d.f. |
| Preferences for News: | |
| General Population | $X^2 = 13.95^*$ 2.68d.f. |
| Low Income White | $X^2 = 0.31$, 3.27d.f. |
| Low Income Black | $X^2 = 0.11$, 2.01d.f. |
9. A more extended exposition of the Coleman analysis applied to a marketing problem is given in Whittaker (1972); see also Alexis, Haines, and Simon (1972).
 10. A second possible explanation is that the world has changed so as to greatly increase random error. While there is no way to exclude this interpretation on statistical grounds, there is also no reason to suppose it is correct.
 11. A "small sample" analysis could be performed in four of the eight cases. The results were not greatly different. If anything, they tended to indicate the importance of the race variable may be slightly overstated in Table X. Copies of these results are available from the authors on request.

References

- Alexis, Marcus, Haines, George H. & Simon, Leonard S. Neighborhood effects, family decision-making patterns, and consumption expenditures. In Fred C. Allvine (Ed.), Combined Proceedings 1971 Spring and Fall Conferences. Chicago: American Marketing Association, 1972.
- Bauer, Raymond, et.al. The marketing dilemma of Negroes. Journal of Marketing, 1965, 29, 1-6.
- Bauer, Raymond A. & Cunningham, Scott M. Studies in the Negro Market. Cambridge, Mass.: Marketing Science Institute, May, 1970.
- Becker, Gary S. A theory of the allocation of time. The Economic Journal. 1965, 75, 493-517.
- Carey, James W. Variations in Negro/White television preferences. Journal of Broadcasting, 1966, 10, 199-211.
- Coleman, James. Introduction to Mathematical Sociology. New York: The Free Press, 1964.
- Fama, Eugene F. Efficient capital markets: a review of theory and empirical work. Working paper, University of Chicago, January, 1970.
- Frank, Ronald E. & Massy, William F. Effects of short-term promotional strategy

in selected market segments. In Patrick J. Robinson (Ed.) Promotional Decisions Using Mathematical Models. Boston, Mass.: Allyn and Bacon, Inc. 1967.

Gerson, Walter M. Mass media socialization behavior: Negro-white differences. Social Forces, 1966, 45, 40-50.

Greenberg, Bradley S. & Dervin, Brenda. Mass communication among the urban poor. Public Opinion Quarterly, 1970, 224-235.

Greenberg, Bradley S. & Dervin, Brenda. Use of the Mass Media by the Urban Poor. New York: Praeger, 1970 (esp. Ch. 1).

Harrison, Bennett. The intrametropolitan distribution of minority economic welfare. Journal of Regional Science, 1972, 12, 23-44.

Honig, David, Bomens, James, Brooks, James & Hickman, Darwin. Some sources of consumer revolt: the market research project, July-September 1971. Paper presented at the Conference on Consumer Affairs, Rochester, New York, June 9, 1972.

McCombs, Maxwell E. Negro use of television and newspapers for political information, 1952-1964. Journal of Broadcasting, 1968, 12, 261-266.

Oladipupo, Raymond O. How Distinct is the Negro Market. New York: Ogilvy and Mather, 1970.

Potthoff, Richard F. & Whittinghill, Maurice. Testing for homogeneity: I. The binomial and multinomial distributions. Biometrika, 1966, 53, 167-182.

Samuelson, Merrill, Carter, Richard F. & Ruggels, Lee. Education, available time, and use of mass media. Journalism Quarterly, 1963, 491-496.

Stafford, James E., Cox, Keith K. & Higginbottom, James B. Some consumption pattern differences between urban Whites and Negroes. Social Science Quarterly, 1968, 49, 614-630.

Westley, Bruce & Severin, Werner J. Some correlates of media credibility. Journalism Quarterly, 1964, 325-335.

Whittaker, William S. The Relationship Between Individual Difference Variables, Media Usage, and Product Choice. Ph.D. Dissertation: University of Rochester, Rochester, N.Y., 1972.

A STUDY OF THE RELATIONSHIP BETWEEN SOCIAL
VALUES AND ATTITUDES TOWARD ADVERTISING

John F. Willenborg¹
University of South Carolina

The study upon which this paper is based investigates the relationship between a person's social values and his attitudes toward the institution of advertising. Employed in the study are three Likert-type scales: a measure of social values or social orientation; an attitude scale made up of a series of statements about various effects of advertising; and a series of descriptions of situations in which the possibility of conflict occurs through the inclusion of both favorable and unfavorable elements. Interrelationships between the scales based on responses from 355 students were measured.

Problem

Consumer research has, for many years, sought to discover more about the dimensions of personality. Motivation, values, and attitudes have been scrutinized from many points of view. This paper reports on a study undertaken to relate social values to attitudes toward advertising and several other business functions.

An important development of recent years has been an increase in pressure being placed on business to respond to the needs of society. It is clear that many institutions of society, including both businesses and certain business functions, no longer can be considered totally autonomous and without social obligations. In response to increasing pressure and, perhaps, with a greater sense of responsibility, they seem to be becoming more sensitive to public and private concern. It follows that research is needed on the dimensions and effect of business practices.

Criticism of the advertising function is not a new phenomenon. However, criticism relating to the marketing function is particularly acute today as consumers, consumer advocates, and government seek to correct abuses and recommend changes.² Although much of the concern is directed toward specific examples of deception, unfair practices, and the like, criticism is often levied against the institution itself. Advertising suffers from an unfavorable image in the minds of many consumers. In this paper, some more light is shed on the nature of the person who tends to criticize, tolerate, or favor advertising and related functions.

It is an underlying assumption of this study that advertising possesses social impact and, perhaps, has become such an important facet of American life that it may have far-reaching social ramifications. Based on this supposition, it is speculated that attitudes and feelings formed by an individual regarding advertising are related to his social values, i.e., social-orientation. Therefore, the major hypothesis tested in this study is:

A person's social value orientation is inversely related to his attitude toward advertising.

Two related hypotheses are:

1. A person's social value orientation is related to the way in which he resolves conflict situations regarding the favorable and unfavorable elements of business practice;
2. Resolution of conflict situations involving favorable and unfavorable elements of business practice is related to a person's attitudes toward advertising.

Research Design

The Measuring Instrument

A questionnaire was developed which is composed of three sections. The first part is a series of twenty statements concerned with various effects of advertising upon society. The statements were selected from thirty statements which had originally been developed. Positive and negative valences for each statement based on overall favorability or unfavorability toward advertising were determined through pre-tests in which valences were assigned by over 100 students. Of the statements selected, ten have positive and ten have negative valences.

Attitudes of respondents toward each statement were recorded on a seven-point Likert-type scale ranging from "strongly agree" to "strongly disagree." Respondents were requested not to proceed to the second part of the questionnaire until additional instructions were given.

The second section includes descriptions of a series of "situations" related to business practices and their effects. The descriptions are designed to introduce conflict through presentation of both positive and negative elements or effects in a marketing and/or business framework. Several statements are directly concerned with advertising and its effects so that their relationship with the other sections of the questionnaire can be tested. Generally, the situations were developed based on frequently-mentioned arguments for and against the practice. The ten statements are presented in Exhibit 1.

Exhibit 1

Descriptions of Conflict Situations*

-
1. Cultural television programs sponsored by companies which also present misleading advertising messages.
 2. New, improved products accompanied by higher prices due to the high cost of product research.
 3. Advertising which exaggerates product benefits, but informs us about what is available to purchase.
 4. Companies which contribute to local charities while their factories pollute the atmosphere.
 5. Steadily increasing salaries for workers accompanied by a higher cost of living.
 6. Advertising which tells people where to buy beautiful, expensive products and often results in buyers going into debt to buy them.
 7. The development and use of new, unbreakable plastic bottles for household cleaners which cannot be "recycled" or easily disposed of when empty.
 8. Advertising on television which is often entertaining, but often seems to insult our intelligence.
 9. The use of new safety devices in automobiles which will make automobile prices higher.
 10. The growth of large chain stores which provide products at lower prices, but force small stores out of business.

*Evaluated by respondents on a scale from "very favorable" to "very unfavorable."

Respondents evaluated each situation in terms of "its overall effect on the American society" on a seven-point scale ranging from "very favorable" to "very unfavorable." A response which tends to be favorable reflects the relative desirability of the positive element and a general willingness of the respondent to accept the negative factor in light of its favorable aspect.

In the third part of the questionnaire, an index of a person's social value orientation was derived through the use of a social value scale developed by S. I. Perloe.³ The scale which is employed consists of twenty statements which are evaluated in order to measure social orientation. More specifically, the statements relate to a person's concern for the welfare of others as a matter of individual preference rather than out of moral obligation. The positive and negative valences assigned by Perloe are utilized⁴ (ten statements have positive and ten have negative valences). Respondents were asked to react to each statement on a seven-point scale on the basis of "strongly agree" to "strongly disagree." Overall social orientation is the mean of all responses.

The Sample and Administration of the Study

The sample was not randomly selected from a large population. However, respondents were chosen in order to fill quotas which were representative of the student populations under study. Three general groups were included: high school students (senior level); lower division college-students (freshman and sophomore levels); and upper division college students (junior and senior levels). Two universities and three high schools were employed for drawing the sample. Total sample size was 355 students.

Administration of the questionnaire took place in small classroom group situations. Average class size was twenty students. Class sizes ranged from seven to forty-five. Administrators clearly defined terms prior to beginning the administration and used a standardized format. Less than one percent of respondent questionnaires were rejected.

Demographic characteristics of respondents are shown in Exhibit 2.

Exhibit 2

Demographic Classifications of Respondents

Education Level		
High school		113
College--lower division		87
College--upper division		147
Race		
Black		91
White		264
Curriculum (college students)		
Business administration emphasis		124
Other (no business administration)		110
Sex		
Male		209
Female		146
Age		
Under 18		57
18-21		231
Over 21		67

Analysis

The major analysis task was to determine the relationships between the three scales: attitudes toward advertising; resolution of conflict situations; and social value orientation. For each respondent, indices were developed by calculating mean scores of responses for each scale. Relationships were studied by computing simple correlation coefficients which indicated the direction and degree of relationship between the three scales for each demographic classification. Significance was measured by application of the t-test. The coefficient of determination was employed to compute the amount of variance explained.

To further understand the relationships, mean scores for each scale by demographic category were analyzed. Finally, responses to the conflict situations were further analyzed by considering the statements as two distinct categories: 1) those dealing directly with advertising; and 2) those involving general marketing and/or business situations.

Results

An inverse relationship between social value orientation and attitude toward advertising had been hypothesized. The correlations lend support to the supposition; i.e., the greater a person's social orientation, the more critical he tends to be regarding advertising. Correlations and levels of statistical significance are presented in Exhibit 3. Generally, the coefficients, while showing the inverse relationship, are not high enough to give clear-cut evidence of the hypothesized relationship. Most are significant at the .05 level (t-test); however, the large number of observations is certainly a contributing factor. Yet, the highest correlations generated in the study relate to the social values-attitude toward advertising relationship. The greatest amount of variance explained by any correlation is nearly twenty percent (coefficient of determination).

No clear pattern was established in analysis of the relationship between social orientation and the resolution of the wide-ranging conflict situations. The hypothesized inverse relationship was supported; i.e., the greater the degree of a person's social orientation, the less likely he is to accept a situation which embodies both positive and negative elements. Likewise, individuals with a lesser degree of social orientation tend to be more favorable to the situations; i.e., they seem willing to accept the negative result as long as the situation contains a compensating favorable factor. Correlation coefficients are generally statistically significant, but are also relatively low for most categories.

It is important to note, however, that correlations are higher in every case when the set of conflict statements specifically involving the institution of advertising are considered as distinct from those concerned with the broader business issues (See Exhibit 3). Although the improvements are not substantial, it appears that the tendency for socially-oriented persons to react unfavorably toward business-related conflict situations is even more pronounced where advertising is concerned. Obviously, the nature of each conflict statement itself may help to account for the result; however, the assumption of the social nature of advertising is given some additional support by the finding.

As expected, a positive relationship exists between attitudes toward advertising and reaction to the conflict situations (See Exhibit 4). A person is generally in favor of the situation described if he is also favorable in his overall view toward advertising. This is particularly true when only the advertising statements are considered.

Exhibit 3

Correlations Between Social Values and Attitudes Toward Advertising and Business Situations by Demographic Category*

Category	Social Values/ Advertising Attitudes	Social Values/ All Situations	Social Values/ Advertising Situations
All respondents (n=355)	-.238 ^b	-.205 ^b	-.244 ^b
Education level			
High school (n=113)	-.181 ^a	-.179	-.235 ^a
Lower division (College) (n=87)	-.430 ^b	-.246 ^a	-.247 ^a
Upper division (College) (n=147)	-.258 ^b	-.225 ^b	-.251 ^b
Race			
Black (n=91)	-.097	-.028	-.104
White (n=264)	-.336 ^b	-.182 ^b	-.240 ^b
Curriculum			
Business administration emphasis (n=124)	-.212 ^a	-.073	-.127
Other (no business administration) (n=110)	-.391 ^b	-.336 ^b	-.351 ^b
Sex			
Male (n=209)	-.219 ^b	-.121	-.205 ^b
Female (n=146)	-.263 ^b	-.317 ^b	-.397 ^b
Age			
Under 18 (n=57)	-.130	-.080	-.103
18-21 (n=231)	-.323 ^b	-.299 ^b	-.323 ^b
Over 21 (n=67)	-.148	-.133	-.202

*Pearson r correlation coefficients based on mean scale values.

^aSignificant at .05 level (t-test).

^bSignificant at .01 level (t-test).

Exhibit 4

Correlations Between Attitudes Toward Advertising and Resolution of Business Conflict Situations By Demographic Category*

Category	Advertising Attitudes/ All Situations	Advertising Attitudes/ Advertising Situations
All respondents	.242 ^b	.285 ^b
Education level		
High school	.185 ^a	.246 ^b
Lower division (College)	.252 ^a	.310 ^b
Upper division (College)	.290 ^b	.314 ^b
Race		
Black	.115	.187
White	.277 ^b	.332 ^b
Curriculum		
Business administration emphasis	.199 ^a	.196 ^a
Other (no business administration)	.240 ^a	.372 ^b

Sex		
Male	.229 ^b	.289 ^b
Female	.267 ^b	.375 ^b
Age		
Under 18	.263 ^a	.301 ^a
18-21	.258 ^b	.317 ^b
Over 21	.190	.184

*Pearson *r* correlation coefficients based on mean scale values.

^aSignificant at .05 level (t-test).

^bSignificant at .01 level (t-test).

When correlations are analyzed by demographic classification, higher correlations are found with college students than with high school students; and higher correlations with white than with black students. Other patterns are not so clearly evident.

Some interesting results derive from analysis of the mean scores for each scale by demographic classification. For example, it would be expected that college students who have been enrolled in business administration courses would be more favorable toward advertising as well as to the business situations than those who have not taken such courses. The mean scores bear out the logic of the argument (See Exhibit 5). Interestingly, there is also a significant difference between the two groups in mean social value score; the students without business courses tend to be more socially-oriented.

There is also a significant difference between blacks and whites in mean scores. Blacks are significantly more favorable toward advertising and business, but tend to be less socially-oriented based on the social value scale (See Exhibit 5). Although these findings must be regarded as only tentative, they suggest that the relationships should be further tested.

Conclusions and Implications

The study described in this paper was exploratory in nature. The social values scale employed, although tested and validated in a variety of experiments, had not been utilized within a business and advertising framework. In addition, the use of situation descriptions with both positive and negative elements represents a somewhat unique approach to attitude measurement. Findings of the research should suggest directions for future research.

Results of the study support the hypothesis that social orientation is inversely related to attitudes toward advertising. Also, the more highly socially-oriented, the less likely a person is to favor a situation in which any benefits may be partially offset by negative factors, particularly if advertising is involved. Although correlations derived from the study are not strong enough to establish the relationship firmly, the direction of the relationship is significant.

It appears safe to speculate that advertising and other business functions do have, at least in the minds of student respondents, social connotations. If this is true, it may be disturbing to those concerned with the management and performance of business activity that socially-oriented persons tend to be anti-advertising. Carrying the reasoning further, it may even be speculated that consumers feel that advertising and related functions are not always in the social interest, a notion that is not unique in the age of consumerism.

From a research technique standpoint, the conflict resolution approach to attitude measurement may provide a means of assessing the strength of one's attitude in the face of potentially offsetting elements. It may be suggested as an alternative approach to attitude measurement which fails to predict

Exhibit 5

Mean Scale Values By Demographic Category*

Category	Scale		
	Advertising Attitudes	Conflict Situations	Social Values
All respondents	3.73	3.73	4.82
Education Level:			
High school	3.75	3.79	4.78
Lower division (College)	3.64	3.61	4.82
Upper division (College)	3.76	3.76	4.84
Race:			
Black	3.84	4.26	4.59
White	3.69	3.56	4.89
Curriculum:**			
Business administration emphasis	3.73	3.87	4.72
Other (no business administration)	3.68	3.53	4.91
Sex:			
Male	3.71	3.77	4.77
Female	3.76	3.67	4.88
Age:			
Under 18	3.74	3.75	4.81
18-21	3.72	3.78	4.81
Over 21	3.76	3.55	4.86

*Mean scores based on 1-7 scale value.

**College students.

behavior because it does not account for the multi-dimensional nature of attitudes or for the presence of such contradictory factors. Obviously, the ultimate test of the measure will lie with further testing and the observance of subsequent behavior.

Footnotes

1. Assistant Professor of Marketing, University of South Carolina.
2. For example, the Federal Trade Commission Hearings on Modern Advertising, November, 1971, dealt with advertising's effects on a broad scale.
3. Described in some detail in John P. Robinson and Phillip R. Shaver, Measures of Social Psychological Attitudes, Institute for Social Research, The University of Michigan, Ann Arbor, Michigan, 1969, pp. 492-501.

The scale is also discussed in a paper by S. I. Perloe in Proceedings of the American Psychological Association Annual Meeting, 1971; and in "Changes in Student's Values and Roles at College," Terminal Progress Report for National Institute of Mental Health Grant, MH 16483-02, by S. I. Perloe.

4. Perloe, in Robinson and Shaver, p. 492.

MATHEMATICAL CONSIDERATIONS IN THE USE
OF LINEAR ATTITUDE MODELS

Frederick W. Winter¹
University of Illinois

Expectancy value models have made a significant impact in the analysis of consumer behavior; the intuitive appeal of the model has led to a broad range of empirical research findings many of which suggest unresolved issues. The basic expectancy value model has been formulated in a variety of ways (Rosenberg, 1956; Fishbein, 1963) and relates the individual's tendency to act as a function of two components:

1. the strength of expectancy that a given act will result in a set of consequences
2. the value of the consequences to an individual

The strength of expectancies are generally multiplied by the value of the consequences and summed over all the salient attributes. It is not the purpose of this paper to review the history of expectancy value models as overall summaries of the research efforts and issues are readily available (Wilkie and Pessemier, 1972; Cohen, 1972). What will be questioned is the operational framework in which consumers' perceptions are translated into the individual's attitudinal measure.

The Basic Concept

In mathematical language, the researcher is asking the individual for his perceptions along some scale, x , and the functional relationship between x and utility or value, $g(x)$. In simple terms one might first ask, "how economical is the Volkswagen?", and "what is your utility for the level of economy offered by the Volkswagen?". Undoubtedly more than one attribute is salient to the individual, and, thus, x_i represents the perception along the i th dimension and the "utility" function $g_i(x_i)$ expresses the relation between value and x_i .

In the majority of research applications x_i is expressed better as a continuous variable as opposed to a dichotomous variable, consequently considering Volkswagen economy over many outcome or consequence states is probably more valid than the two-state representation of "economical" or "uneconomical." As with most perceptual processes, it is unlikely that an individual's perceptions along one attribute can be summarized with a point estimate. Uncertainty is one example where a probability distribution is needed to adequately describe the perceptual framework.

Expansion of the Basic Concept

Figure 1 can serve as an example of a consumer's perception of Volkswagen economy. Researchers have, in the past, asked for one of two types of responses:

- 1) "on a 1 to 7 scale, how economical is the Volkswagen?"
- 2) "how improbable (1) - probable (7) is it that Volkswagen is economical?"

The first question presumably asks for a modal or mean estimate from the individual's perceptual distribution; this is designated as $x_{i,m}$ in Figure 1. To

respond to the second question the individual must first designate a boundary point that divides the economy continuum into the uneconomical-economical dichotomy. This boundary point is designated as $x_{i,b}$, and the question asks for the probability that Volkswagen is greater than $x_{i,b}$ [i.e., $\sum_{j=b}^{\infty} f(x_{i,j})$]. Whether or not consumers' responses correspond to their distribution is a question to be answered by future research efforts.

The second component of the expectancy value model is a functional form of $x_{i,j}$ to relate consequences to utility $g_i(x_{i,j})$. A monotonic relation between utility and perception is shown in Figure 2. Clearly it is possible to imagine a variety of circumstances where the relation would not be monotonic (e.g., temperature or sweetness of coffee, power in an automobile, etc.)

Utilizing the information provided by question 1, the estimate of VW utility if a modal response is assumed is:

$$(1a) \quad U_{VW} = \sum_{i=1}^n g_i(x_{i,m}) = \sum_{i=1}^n g_i [x_{ij} \text{ at Max } \{f(x_{i,j})\}]$$

where n = total number of salient attributes.

A mean response results in:

$$(1b) \quad U_{VW} = \sum_{i=1}^n g_i(x_{i,m}) = \sum_{i=1}^n g_i [\sum(x_{i,j}) f(x_{i,j})]$$

Question 2 responses yield:

$$(2) \quad U_{VW} = (\text{Expected Utility of } x_{ij} \text{ where } j > b) (\text{probability } x_i > x_{i,b})$$

$$= \sum_{i=1}^n \left[\frac{\sum_{j=b}^{\infty} g_i(x_{i,j}) \frac{f(x_{i,j})}{\sum_{j=b}^{\infty} f(x_{i,j})}}{\sum_{j=b}^{\infty} f(x_{i,j})} \right]$$

$$= \sum_{i=1}^n \sum_{j=b}^{\infty} g_i(x_{i,j}) [f(x_{i,j})]$$

An alternative measure might be to consider an expected utility model that considers probability multiplied by utilities over all possible attribute values. A model of this nature is currently undergoing testing (Ahtola, 1971). Using current notation:

$$(3) \quad U_{VW} = \sum_{i=1}^n \sum_{j=0}^{\infty} g_i(x_{i,j}) [f(x_{i,j})]$$

Although it is impossible to specify (1), (2), or (3) as being more desirable at this time, it is obvious that different perceptual measures will result in different attitude measures. Equation (3) does appear to be the more appealing measure since (1) and (2) eliminate some of the information. It can be seen that (1a) and (3) will yield comparable results only when:

$$g_i(x_{i,j}) = \text{constant, for all } j$$

$$\text{or, } f(x_{i,j}) = 0 \text{ when } j \neq m \quad (\text{total certainty})$$

$$= 1 \text{ when } j = m$$

or some less than obvious sets of $f(x_{i,j})$, $g_i(x_{i,j})$.

The same conditions result in identical (1b) and (3) measures with the additional possibility when the utility function is linear:

$$\text{if } g_i(x_{i,j}) = a + b(x_{i,j})$$

$$(3) U_{VW} = \sum_{i=1}^n \sum_{j=0}^{\infty} [a + b(x_{i,j})] f(x_{i,j})$$

$$= \sum_{i=1}^n [\text{const} + b \sum_{j=0}^{\infty} (x_{i,j}) f(x_{i,j})]$$

$$= \sum_{i=1}^n [\text{const} + b(x_{i,m})] = \text{Equation (1b)}$$

Equations (2) and (3) yield similar measures when:

$$\sum_{j=0}^b f(x_{i,j}) = 0$$

$$\text{or, } g_i(x_{i,j}) = 0 \text{ for all } j < b$$

or, some less than obvious sets of $f(x_{i,j})$, $g_i(x_{i,j})$

Equation (3) seems to correspond best to the spirit of the expectancy-value concept by considering probabilities and utility over all possible ranges of $x_{i,j}$ values; only in selected cases will (1a), (1b), and (2) result in identical measures to (3). It will be important in future research efforts to not only empirically compare the alternative models, but to also determine why different results occurred. A natural starting point is to determine how consumers respond to alternative questions and how these responses relate to the perceptual distribution of Figure 1.

The lack of correspondence between (1a), (1b), (2), and (3) is particularly evident in the study of attitude change. Figure 3 illustrates what might be pre-post measures of the perceptual space. In this example, mean or modal [(1a) and (1b)] perception measures have not changed. Boundary probability measures [(2)] reflect change that is due to uncertainty reduction alone. Equation (3), probability measures over all possible consequences, utilize the entire information that mean responses haven't changed while uncertainty measures have.

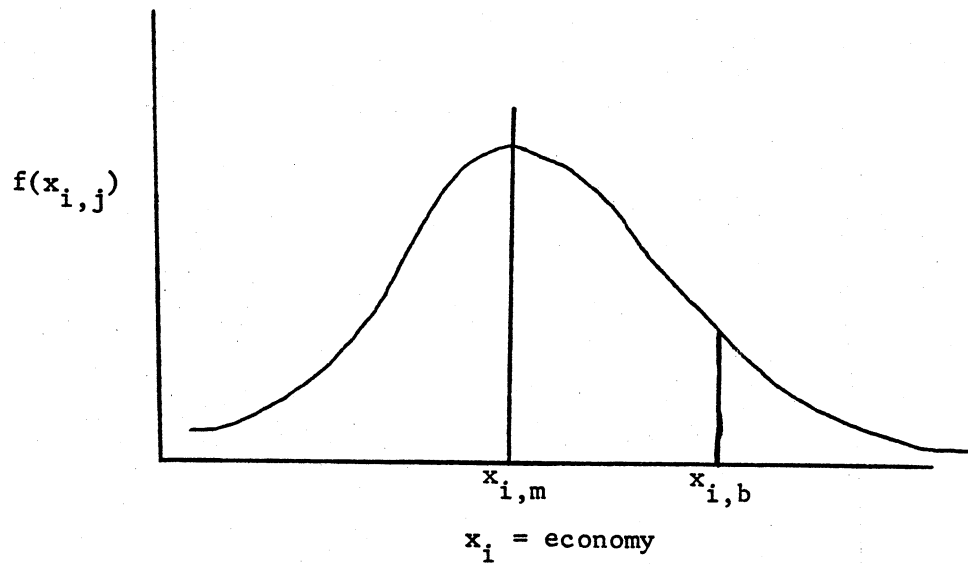
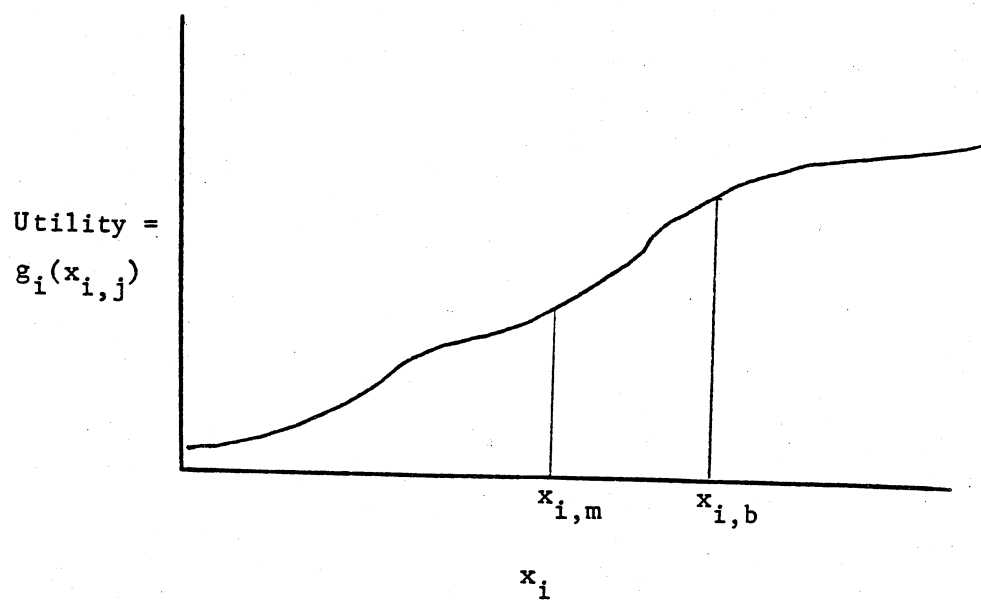


FIGURE 1. Perceptual Representation of VW Economy

FIGURE 2. Utility Function of $x_{i,j}$

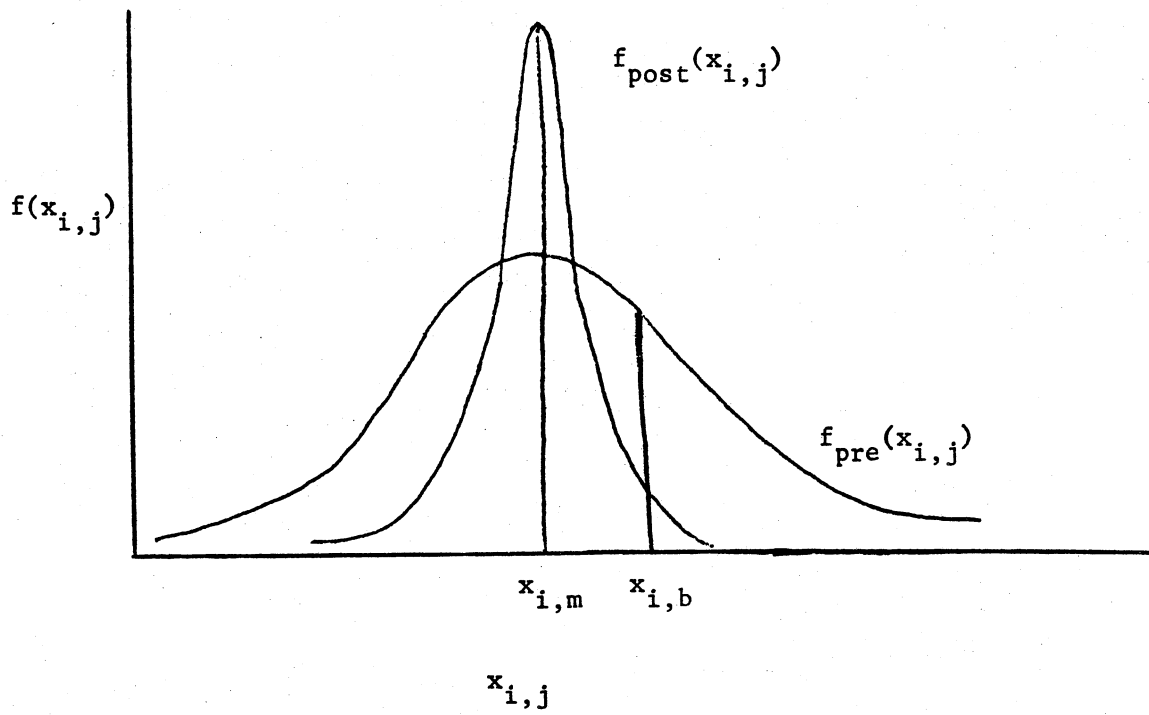


FIGURE 3. Pre-Post Measures of the Perceptual Space

Summary

There is no one correct way to measure perceptions and utility-values in expectancy value models. It is important, however, to be aware that different measures provide different information about potential consumer behavior. Hopefully, this view of the measures will stimulate empirical research to consider not only which measures may be applicable but also to understand why the measures may differ.

Footnotes

1. Frederick W. Winter is Assistant Professor of Business Administration, University of Illinois at Urbana-Champaign.

References

- Ahtola, Olli T. "An Investigation of Cognitive Structure Within Expectancy-Value Response Model," dissertation proposal presented at College of Commerce and Business Administration, University of Illinois, 1971.
- Fishbein, Martin. "An Investigation of the Relationships Between Beliefs About an Object and the Attitude Toward that Object," Human Relations, 1963, 16, 233-40.
- Rosenberg, Milton, J. "Cognitive Structure and Attitudinal Affect," Journal of Abnormal and Social Psychology, 1956, 53, 367-72.
- Wilkie, William L. and Edgar A. Pessemier. "Issues in Marketing's Use of Multi-Attribute Attitude Models," Paper No. 365, Institute for Research in the Behavioral, Economic, and Management Sciences, Krannert School of Industrial Administration, Purdue University (August, 1972).

ATTITUDES AS PREDICTOR OF WHAT?: EXPLORING THE STRUCTURE
OF PRODUCT-RELATED BEHAVIORS

Alain Cousineau, University of Sherbrooke¹
Peter L. Wright, University of Illinois

Perhaps no other concept has generated more theorizing and experimentation in the history of social psychology than the concept of attitude (Allport, 1935, 1954) and it is not surprising that it came to play an increasingly important role in most of the behavioral sciences and the applied disciplines, marketing and consumer research being no exception. The popularity of the attitude construct among behavioral scientists and the consumer research community is probably best explained by the basic assumption underlying most of its proposed definitions: an attitude, being essentially a learned predisposition to respond toward an object in a consistently favorable or unfavorable manner (Allport, 1935), should be functionally related to behavior and, to that extent, should provide a sound theoretical basis for explaining and predicting that behavior.

However, as was the case for social psychologists (LaPierre, 1934; Bray, 1950; Kutner, Wilkins and Yarrow, 1952; Berg, 1966; etc.), consumer researchers have often found little evidence of a strong attitude-behavior relationship. Fishbein (1967) has pointed out that there are at least two reasons for these poor results. First, attitudes are often measured with respect to a general class of objects rather than a specific object. Second, the behavior predicted may be only partially related to attitude. That is, various responses may be learned for any given attitude so that two persons may learn to hold the same attitude toward a given stimulus but they may also learn to make different responses given the learned attitude (Doob, 1947). Because of these considerations, Fishbein (1967) has argued that if prediction and explanation of behavior is the focus of interest, then the orientation taken to the concept of attitude and the means of measurement and analysis used must be consistent with that focus. Accordingly, Fishbein (1967) has proposed to measure attitudes toward specific actions rather than toward the objects which provide the occasion for behaving. Ajzen and Fishbein (1970) have provided evidence to support the notions that these two types of attitude are not necessarily related and that attitudes toward specific actions are better predictors of behavioral intentions.

However appealing this orientation to the concept of attitude may be from a theoretical perspective, a number of objections can be raised against it. While consumer researchers are mostly interested in predicting and explaining purchase behavior, they are also concerned with a host of other product-related behaviors (e.g., communication behaviors, usage behaviors). According to the suggested orientation, no single index of attitude will provide a valid basis for predicting all these behaviors of interest since an attitude toward a given product-related behavior to be performed in a specific context (e.g., private use of a product) is not expected to be a good predictor of the same behavior occurring in a different context (e.g., public use of the product) or of a different product-related behavior (e.g., giving advice to a friend about the product). Furthermore, the set of salient consequences perceived by a consumer to be the likely results of the performance of a given product-related behavior --

which is assumed to be a major determinant of his attitude toward that behavior -- will probably not remain invariant across behaviors and contexts. If it does, the consequences are expected to carry different weights in determining the various attitudes. From these considerations it follows that if attitudinal data of the kinds suggested by Fishbein (1967) are to be used to predict and explain how consumers behave globally toward a given product or brand, researchers are faced, at least theoretically, with the impractical task of having to obtain measures of attitudes toward each possible combination of behaviors and contexts involving the product or brand, and having to define, for each of these combinations, a set of salient consequences.

These objections do not invalidate the notion of using attitudes toward specific actions as a means of predicting and explaining specific product-related behaviors occurring within specific contexts. Rather they point out to some of the problems involved in trying to predict and explain how consumers behave in a more global fashion toward products since no single index of attitude toward an act is likely to be very useful in considering all possible behaviors and contexts involving a given product. Clearly means of coping with this problem must be sought. Perhaps one way of dealing with this issue is to look more closely at the dependent variable, that is, the behavior with respect to a product. Much of the research emphasis until now has focused on issues related to the structure of attitudes (Pessemier and Wilke, 1972). Little research, if at all, has been done on what we mean by "behavior with respect to a product," or, more importantly, what it means psychologically for the consumer. Clearly, by understanding better this construct -- behavior with respect to a product -- we will be able to solve some of the problems raised earlier.

For instance, it may be the case that not all the behaviors that one could engage in with respect to a given product are psychologically "pure" or independent of each other in the mind of the consumer, thus suggesting the possibility of identifying those behaviors or groups of behaviors which are perceived as psychologically equivalent. The identification of such clusters or groups of behaviors, each one presumably reflecting a more fundamental dimension of behavior, would then make it unnecessary to specify all the possible behaviors that one could engage in with respect to a product and would allow us to limit the analysis to a (hopefully) small subset of more fundamental dimensions of behavior.

It may also be the case, as was suggested by Cohen and Ahtola (1971) that a small number of situational contexts might be specified in advance for each of these fundamental dimensions of behavior which could then be incorporated into the measurement procedure and used in prediction after the assignment of appropriate probabilities of occurrence.

An exploratory investigation, dealing with these issues, is currently under way. Essentially, this investigation seeks to uncover the factor structure underlying the construct "behavior with respect to a product," that is, it explores the possibility of identifying a few independent behavioral tendencies that could adequately describe this construct.

The methodology used is essentially similar to the one employed by Triandis (1964) to explore the factor structure of the behavioral components of social attitudes. The variables used in this exploratory study create a matrix of data. One face of this matrix consists of a set of 16 hypothetical products that are described solely in terms of four attributes: conspiruosity, relatedness to self-image, familiarity and

expensiveness. The second face of the matrix consists of 30 behaviors that a consumer might undertake in relations to such stimulus products, for example, purchase them on a trial basis, seek information from a salesman about them, ask for them as a gift from a good friend. A typical judgment made by a subject is exemplified by the following item:

conspicuous, very familiar, inexpensive, closely related to my self-image
 would:____:____:____:____:____:____:____:____:____:would not
 make a special trip to purchase it

The matrix of data thus generated will then be analyzed through factor analytic procedures to determine if a simple structure can be identified that could adequately describe the construct "behavior with respect to a product."

The isolation of dimensions underlying the construct "behavior with respect to a product" has several advantages. It permits clearer analysis and, hence, better understanding of the theoretical construct. It permits to identify explicitly the dimensions underlying the construct, through a few behavioral tendencies that are independent of each other. Finally, it may help us solve some of the problems involved in using attitudes toward specifications as a means of predicting and explaining how consumers behave in a more global fashion toward products and brands.

¹Alain Cousineau is on the faculty at Sherbrooke University (Canada) and is presently on leave to complete doctoral work at the University of Illinois. Peter L. Wright is Assistant Professor, Department of Business Administration, University of Illinois.

References

- Ajzen, I. & Fishbein, M. Attitudinal and normative variables as predictors of specific behaviors: a review of research generated by a theoretical model. Paper presented at the Workshop on Attitude Research and Consumer Behavior, University of Illinois, December, 1970.
- Allport, G. W. Attitudes. In C. Murehison (ed.), A handbook of social psychology. Worcester, Mass.: Clark University Press, 1935.
- Allport, G. W. The historical background to modern social psychology. In G. Lindzey (ed.), Handbook of social psychology, Vol. 1, Theory and method. Cambridge, Mass.: Addison-Wesley, 1954.
- Berg, K. E. Ethnic attitudes and agreement with a negro person. Journal of Personality and Social Psychology, 1966, 4, 215-220.
- Bray, D. W. The prediction of behavior from two attitudes scales. Journal of Abnormal and Social Psychology, 1950, 45, 64-84.
- Cohen, J. B. & Ahtola, O. T. An expectancy x value analysis of the relationship between consumer attitudes and behavior. In D. M. Gardner (ed.), Proceedings, Association for Consumer Research, 1971, 344-362.
- Doob, L. The behavior of attitudes. Psychological Review, 54, 135-156.

- Fishbein, M. Attitude and the prediction of behavior. In M. Fishbein (ed.), Readings in attitude theory and measurement. New York: Wiley, 1967, 477-492.
- Kutner, B., Wilkins, C., and Yarrow, P. R. Verbal attitudes and overt behavior involving racial prejudice. Journal of Abnormal and Social Psychology, 1952, 47, 649-652.
- LaPierre, R. T. Attitudes vs. actions. Social Forces, 1934, 13, 230-237.
- Pessemier, E. A. and Wilke, W. L. Multi-attribute choice theory - a review and analysis. Institute Paper No. 372, Institute for Research in the Behavioral, Economic and Management Sciences, Purdue University. September, 1972.
- Triandis, H. C. Exploratory factor analyses of the behavioral component of social attitudes. Journal of Abnormal and Social Psychology, 1964, 68, 420-430.

INVESTIGATING THE FEASIBILITY OF PERSONALIZED

RAPID TRANSIT: AN EXPERIMENTAL APPROACH

Richard J. Lutz
University of Illinois

One of the most serious challenges of the 1970's is that of cleaning up the air in major U. S. cities. At the present time, the more than 87-million automobiles on the road are considered to be the cities' worst single source of air pollutants.

For years the primary alternative to the automobile in most U. S. cities has been antiquated and highly inefficient public transportation. However, new technology now makes it possible to automate public transportation, thus greatly reducing operating costs. A revamping of public transportation modes will do much to help clean up American cities.

One of the more exciting new developments in public transportation is the concept of "personalized rapid transit," which provides an alternative to the automobile for short-haul service within major metropolitan areas. Personalized rapid transit (PRT) calls for a network of computer-controlled cars, each of which carries 20 or fewer passengers. During non-peak load times, a passenger would be able to use the PRT vehicle to reach a specific destination directly--the vehicle would not follow a pre-designated route. During rush hours, the PRT system could be adapted to operate like a regular mass transit system, with trains of vehicles following specific schedules.

One of the problems which has faced designers developing PRT vehicles has been the size of the vehicle. While a small vehicle would be desirable, designers fear the reaction of passengers to being enclosed, perhaps with strangers, in a vehicle of the dimensions of the automobile.

This vehicle size problem is especially evident in the design of a PRT system for an existing tunnel network in a large Midwestern city. Shortly after the turn of the century, a tunnel network was completed which was used for freight delivery to major buildings in the city's downtown district. The company that operated the system went bankrupt in 1959, and since that time the city has been searching for a new use for the tunnels.

One recent proposal which is being considered by the city is to develop the tunnel network into a PRT system. As one segment of the feasibility studies being conducted regarding this proposal, the Research and Design Department of the University of Illinois developed a mock-up of the proposed PRT vehicle.

A unique problem which was encountered in the design of the vehicle resulted from the dimensions of the tunnels for which it was being designed. With a height of only 7'6", the tunnel would allow a vehicle of no more than 5' interior height. This would mean that passengers in a moving vehicle would be unable to stand. Feeling that this feature might be a strong deterrent to potential passengers, it was decided to consider an alternative vehicle which would allow passengers to stand if they desired. If this alternative vehicle were found to offer a significant advantage over the smaller vehicle, the tunnels would have to be enlarged by some two feet in height, at a cost of several million dollars. Thus, an important decision in the design of the PRT vehicle are passengers' reactions to its size. The primary purpose of this investigation,

therefore, was to determine the effects of vehicle size on passenger attitudes toward a mock-up of the proposed PRT vehicle.

A secondary concern was the load factor--or how the presence of other passengers affects passenger reactions to the vehicle. Recent research on "crowding" (Little, 1965; Griff and Veitch, 1971; Stokols, et al., 1972) has shown that people tend to react negatively toward situations in which their "personal space" is limited. It was felt that a passenger in a fully loaded PRT vehicle might attach a lower evaluation to the vehicle than would a single passenger.

Theoretical Orientation

A general class of attitude models--labeled "expectancy-value" models--has seen increasing use in marketing in the past few years. These types of models are most often direct applications or hybrids of models developed by Fishbein (1963) and Rosenberg (1956), although some have been developed within marketing (Cohen and Houston, 1970; Sheth, 1970). Essentially, these models are all based on the premise that a person's attitude toward any object, action, or brand can be represented by his cognitions or beliefs about the attitude object. While some issues remain unresolved as to the appropriate measures and combination rules implied by these models, the most general form of the model can be represented algebraically:

$$A_o = \sum B_i a_i$$

where A_o = attitude (affect) toward the object, action, or brand under investigation

B_i = the belief that the object possesses some attribute i or will help reach some goal i

a_i = attitude toward or satisfaction with the i^{th} attribute or goal.

The expectancy-value model was chosen as the theoretical orientation which guided the investigation of passenger reactions to the proposed PRT vehicle. In addition to offering insights into the relative merits of the two alternative vehicles, it was hoped that such a test would provide more evidence regarding the usefulness of the expectancy-value model as a decision-making tool. If it was found that passengers did develop a cognitive structure regarding the PRT vehicle after a short exposure to it, then there would be more justification for believing that consumers really do form attitudes on the basis of their beliefs about products and brands.

Four hypotheses were tested in this study:

1. There is a positive relationship between attitude and cognitive structure as a result of direct experience with the PRT vehicle.
2. There is a significant interaction between vehicle size and the load factor in the PRT vehicle. Specifically, the large vehicle is rated relatively higher than the small vehicle in the single passenger condition; and due to the crowding factor, the small vehicle is rated relatively higher than the large vehicle in the fully loaded condition.
3. Differences in attitude toward the PRT vehicle across experimental conditions are reflected in corresponding differences in the index of cognitive structure ($\sum B_i a_i$).

4. Differences in attitude across experimental conditions are reflected in differences in single B_i and a_i components of cognitive structure.

Subjects

Subjects were 65 undergraduate students enrolled in introductory marketing, management, and psychology courses at the University of Illinois. Since some of the experimental conditions involved running subjects in small groups, effort was taken to prevent groups of friends from comprising an experimental group. It was felt that being in the vehicle with a close friend would tend to bias the Subject's responses. All Subjects were volunteers who agreed to participate after hearing a short description of the project. At no time was the study referred to as an experiment--rather, it was called a "marketing research study."

Apparatus

The main piece of apparatus used in this study was the mock-up of the proposed PRT vehicle. This mock-up had previously been constructed by Mr. Paul Eshelman, who assisted with this study. The vehicle was adjustable to two sizes, which comprised one factor under investigation here. The "small" vehicle was 5' high and 7' long with four seats facing each other. It could accommodate no standing passengers, so the fully loaded condition consisted of four seated passengers. The "large" vehicle was 7' high and 9' long, again with four seats. Since there was now room for standing passengers, the fully loaded condition consisted of four seated and four standing passengers. Thus, when fully loaded, the large vehicle was actually more "crowded" than the small vehicle since twice as many people were in the only slightly larger vehicle. While the mock-up was stationary, it was windowless; this latter feature made it possible to create a sensation of movement during the simulated "trip" in the vehicle. A small electric motor was placed against the side of the vehicle, producing sound and slight vibrations. This measure was not intended to be a deception, as it was obvious to Subjects that the mock-up was immobile, but pretests had shown that the use of the motor did tend to add a degree of realism to the "ride."

Experimental Situation

The Subject(s) entered a large room and sat down at a long table. After introducing himself, the Experimenter explained the procedure which would be followed in the study. This consisted of three parts:

1. A 2-minute slide presentation providing information about PRT systems in general.
2. A 2-minute "ride" in the PRT vehicle.
3. Completing a short questionnaire.

The Subjects then saw the slide presentation, which was accompanied by a tape-recorded message.

Immediately following the slide presentation the Subject was read some additional information regarding the tunnel network for which the PRT vehicle was being designed. Following the reading of the tunnel information, the Subject

was led into another room which contained the mock-up of the PRT vehicle. He was instructed to take a seat in the vehicle and not to talk (if in a group condition) with the other "passengers." Using a stopwatch to time the "trip," the Subject was left in the vehicle for exactly two minutes.

Upon completion of the trip, the doors of the vehicle were opened and the Subject was ushered back into the main room. At this point, he completed a questionnaire which tapped his cognitive structure regarding the PRT vehicle as well as obtaining several measures of attitude toward the vehicle and the slide presentation. After the Subject had completed the questionnaire, the Experimenter explained the nature of the study and cautioned the Subject not to reveal the purpose of the study to his classmates who had not yet participated. After answering any further questions, the Experimenter thanked the Subject for his cooperation and terminated the experiment.

Analysis and Results

Subjects in the small vehicle--single passenger and large vehicle--single passenger conditions were all run individually. Subjects in the small vehicle--fully loaded condition were run in groups of four, and Subjects in the large vehicle--fully loaded condition were run in groups of eight. Due to Subjects not showing up when scheduled, two graduate students served as "stooges" to ensure a fully loaded vehicle. Since the smaller vehicle allowed no standing passengers, the standing--seated factor could not be investigated across the two sizes of vehicles. The only meaningful analysis involving the standing passengers in the large vehicle--fully loaded condition would be a t-test of the difference between the means of that group and the seated group in that same condition.

Attitude toward the PRT vehicle was assessed in three ways:

1. A semantic differential consisting of 15 scales. A principal components analysis yielded three scales which loaded highly on the evaluative factor (good-bad; ugly-beautiful; pleasant-unpleasant). The scores on these three scales were summed to form one measure of A_o .
2. A simple 1-10 rating scale of affect toward the vehicle, ranging from "like not at all" to "like as much as possible."
3. The derived $\Sigma B_i a_i$ index. Thirteen attributes of the vehicle were measured which had been identified as salient in open-end pretests. All B_i were measured via Sheth-type bipolar scales ("evaluative beliefs"); e.g., "dimly lit"-"brightly lit." All a_i were measured on 7-point bipolar scales ranging from "good" to "bad;" e.g., "The lighting in the vehicle was..."

The correlation between $\Sigma B_i a_i$ and the semantic differential measure was .70. The $\Sigma B_i a_i$ --rating scale correlation was .53. Thus it appears that the cognitive structure index did provide a good measure of attitude toward the PRT vehicle. The semantic differential--rating scale correlation was .67, so all three attitude measures were strongly interrelated.

In testing the load--size interaction hypothesis, a 2x2 factorial ANOVA was conducted on each of the attitude measures. All three showed a significant main effect for the size factor, but did not show a significant interaction effect, thus invalidating the second hypothesis. Table 1 summarizes the ANOVA results.

TABLE 1

Measure	Means		F	Prob.	$\hat{\omega}^2$
	Small	Large			
Semantic Differential	16.0	13.7	9.68	0.003	.15
$\Sigma B_i a_i$	351.6	266.9	13.33	0.001	.19
Rating Scale	7.6	6.7	3.65	0.06	.05

An unexpected finding was that the smaller vehicle was rated higher than the larger vehicle. Calculation of Hays' (1963) omega-square statistic shows the amount of variance explained by the size factor for each of the attitude measures. From observation of Table 1, it is clear that the cognitive index measures of attitude ($\Sigma B_i a_i$) does behave similarly to the other attitude measures, thereby supporting the third hypothesis. Additionally, the ANOVA using $\Sigma B_i a_i$ as the dependent variable was able to account for more of the variance in passenger reactions to the vehicle than either of the other attitude measures, thus lending further support to the use of the expectancy-value approach in the assessment of consumer attitudes.

Of interest are the differences in specific elements of cognitive structure which lead to the differences in overall attitude toward the PRT vehicle. Of the 13 B_i and 13 a_i , a total of eleven significant (beyond $\alpha=.05$) main effects were obtained from 2×2 factorial ANOVA. Table 2 summarizes these findings. All the main effects for vehicle size were in the same direction as the effects for overall attitude--i.e., the small vehicle. While it is beyond the scope of this report to speculate as to the reasons underlying this pattern of results, it is interesting to note that a_i components were significant in three cases where there was no significant B_i effect for the corresponding B_i components. While this may be a non-predicted

TABLE 2

Attribute	B_i	a_i
Interior Space	-	-
Color	-	*
Lighting Glare	-	-
Shape	-	-
Ventilation	*	*
Temperature	-	-
Entry/Exit	-	-
Lighting Level	-	*
Comfort	-	-
# Seats	*	*
View	*	*
Seat Arrangement	-	*
Interest	*	*

*=Significant effect

result, it is not too surprising. Cohen and Houston (1971) have demonstrated that B_i components can differ, depending only upon usage of a particular brand. There is no reason to suspect that a_i components would not behave similarly in certain situations. If a valued object is seen as possessing certain desirable qualities, the consumer may actually re-evaluate upward his liking for or satisfaction with those qualities. However, the results presented here only suggest such a process. Additional experiments are required to test that hypothesis.

Conclusion

The results of this study are promising, from both a theoretical and a practical point of view. The expectancy-value model proved to be quite useful in determining reactions to the PRT vehicle, and the potential seems great for further applications of the model in the area of public policy decisions.

Secondly, this study illustrates the age-old claim that "there is nothing as practical as a good theory." Not only did this study provide useful information to designers working on PRT vehicles, but it also contributed to the body of literature surrounding theoretical properties of the expectancy-value model.

Obviously, there are limitations involved in this study--the results in no way indicate how potential passengers might respond to the total PRT system. Only reactions to the design of the vehicle were considered. The sample was small and non-random, thus weakening the generalizability of the results. From a theoretical standpoint, this study can best be viewed as exploratory in terms of the hypotheses which can be generated from the results.

The primary contribution of the study was to lend further support to the usefulness of the expectancy value attitude model: it showed that a cognitive structure is developed rather quickly through direct experience with an attitude object. Previous studies have tended to deal with long-established attitudes or attitudes newly formed on the basis of information.

References

- Cohen, J. & Houston, M. Some Alternatives to a Five-Point Likert Scale (Especially if You Have a Purpose in Mind). Paper presented at Attitude Workshop, University of Illinois, December, 1970.
- Cohen, J. & Houston, M. Cognitive Consequences of Brand Loyalty. Journal of Marketing Research, 1972, 9, 97-99.
- Fishbein, M. An Investigation of the Relationships Between Beliefs About an Object and the Attitude toward that Object. Human Relations, 1963, 16, 233-240.
- Griffit, W. & Veitch, R. Hot and Crowded: Influences of Population Density and Temperature on Interpersonal Affective Behavior. Journal of Personality and Social Psychology, 1971, 17, 92-98.
- Hays, W. Statistics for Psychologists. New York: Holt, 1963.
- Little, K. Personal Space. Journal of Experimental Social Psychology, 1965, 1, 237-247.

- Rosenberg, M. Cognitive Structure and Attitudinal Affect. Journal of Abnormal Social Psychology, 1956, 53, 367-372.
- Sheth, J. An Investigation of Relationships Among Evaluative Beliefs, Affect, Behavioral Intention and Behavior. Unpublished working paper, University of Illinois, 1970.
- Stokils, D., Rall, M., Pinner, B., & Schopler, J. Physical, Social, and Personal Determinants of the Perception of Crowding. Unpublished working paper, University of North Carolina, 1972.

BELIEFS ABOUT OTHERS AS DETERMINANTS
OF PURCHASE BEHAVIOR

Robert E. Burnkrant
Illinois State University

Understanding and predicting purchasing and the decision making which underlies it is an objective of interest to everyone in consumer behavior. Much recent work directed at furthering this understanding and ability to predict has centered around expectancy-value models of attitude and intention formation and change with notable emphasis recently being placed on the Fishbein model.

A theoretical approach which offers considerable insight into the processes underlying the formation of attitudes and intentions with respect to products is attribution theory. This theory, although receiving considerable attention in psychology, has received only minor consideration thus far in consumer behavior. Settle (1972) and Settle, Faricy and Warren (1971) have recently done some work in this area applying an approach to attribution theory which draws heavily on the work of Kelley (1967).

Another approach to attribution theory which is being taken by Calder and Burnkrant (in final preparation) is derived primarily from the work of Jones and his associates (Jones and Davis, 1965; Jones and Harris, 1967). This espousal of attribution theory focuses more directly and explicitly on the formation and change of beliefs about other users of products based on observations of their behavior with respect to these products. Questions that might be asked of consumers and are currently being investigated at the University of Illinois are to what extent consumers attribute characteristics to others (acquire beliefs) based on the others' use of certain products or brands, and to what extent consumers are aware of and behave in accordance with these attributions.

In the initial study of a series of research projects currently being conducted to investigate these phenomena an attempt was made to determine if the types of products people use are significant factors in determining the characteristics others attribute to them. Although analysis of this data is not yet complete, preliminary results indicate that people make significantly different ($p < .01$) attributions to users of brands which are low in social desirability than they do to users of brands which are high in social desirability (Calder and Burnkrant, in final preparation). To state this more simply they attribute distinct characteristics to individuals based on their use of brands of a product which vary in social desirability.

The study to be discussed here was undertaken as a follow-up to the initial project to see if these results which were obtained using college students in a relatively structured homogeneous environment could be replicated in a field survey using people taken from the community at large and if the respondent's purchase of a given brand influences the attributions he makes about other users of that product. Specifically, it was hypothesized that users of a product low in social desirability would be perceived as having characteristics significantly different from users of a product high in social desirability. It was also predicted that purchase by the respondent of one of these types of products would not significantly affect the attributions made on the basis of another's use of one of these product types.

TABLE 1

Tests of the Multivariate F Test				
Source	F	DF.Hyp.	DF.ERR.	p less than
Interaction	.526	3	89	0.665
Product Purchased	1.210	3	89	0.311
Woman Evaluated	4.241	3	89	0.008

TABLE 2

Woman Evaluated - Main Effect			
Dependent Variable	F	mean square	p less than
Factor 1	3.797	359.453	0.054
Factor 2	4.275	300.190	0.042
Factor 3	12.339	88.606	0.001

FACTOR 1 :

TABLE 3

		Product Purchased		Marginals
		Regular	Non-phos	
Product Evaluated	Regular	38.6	43.6	39.6
	Non-phos	42.6	46.8	
Marginals		40.8	45.5	43.5

FACTOR 2:

TABLE 4

		Product Purchased		Marginals
		Regular	Non-phos	
Product Evaluated	Regular	33.7	33.9	33.7
	Non-phos	36.3	40.7	
Marginals		35.2	38.0	37.3

FACTOR 3:

TABLE 5

		Product Purchased		Marginals
		Regular	Non-phos	
Product Evaluated	Regular	7.5	7.6	7.5
	Non-phos	5.8	4.9	
Marginals		6.5	6.0	5.6

Methodology

A probability sample of 100 respondents was selected from the Champaign-Urbana area. The Champaign Urbana City Directory (1972) which lists streets and addresses alphabetically and numerically was sampled systematically to provide 20 blocks. Within each selected block a complete listing of housing units was obtained. Five people were systematically selected from within each block and they became the respondents for the sample.

Each selected respondent was personally interviewed by trained student interviewers who provided subjects with a structured questionnaire and written instructions. The critical items for the purpose of this paper were a series of 22 semantic differential scales on which the respondents rated either a woman who "normally uses a regular phosphate detergent in doing her laundry," or a woman who "normally uses a non-phosphate detergent in doing her laundry." The determination of which woman the respondent evaluated was made by placing one description systematically in half the questionnaires and the other description in the other half of the questionnaires. Thus each successive respondent received a different treatment level. Each respondent was also asked to indicate whether she usually bought a regular phosphate detergent, a non-phosphate detergent or didn't know.

Results

The semantic differential scales were factored using principal components analysis on a correlation matrix adjusted for the treatment effects. Three factors were extracted and rotated using a varimax rotation. Those items loading highly on a given factor were selected as defining that factor. The individual's score on that factor became the sum of his scores on the selected scales. The scales selected to represent each factor are listed as follows:

<u>Factor 1</u>	<u>Factor 2</u>	<u>Factor 3</u>
Valuable	Happy	Uninformed
Popular	Wise	Impulsive
Graceful	Feminine	
Talkative	High Status	
Neat	Secure	
Economical	Interesting	
Serious	Successful	
Concerned	Mature	
Attractive		
Liberal		

Using these factor scores as dependent variables a 2 X 2 factorial analysis of variance was performed. The two treatments were whether the woman evaluated by the respondent bought regular detergent (low) or non-phosphate detergent (high) and whether the respondent purchased a regular detergent (low) or a non-phosphate detergent (high). A significant multivariate F test ($p < .01$) was obtained for the main effect of the "woman evaluated" treatment. The multivariate F test obtained for the "product purchased" treatment was not significant. The interaction effect was also tested and found not to be significant.

Looking at the univariate F tests for the three factors on the "woman evaluated" treatment we find the first factor marginally significant at $p < .054$, the second factor significant at $p < .05$ and the third factor significant at $p < .001$.

Discussion

A look at Tables 3, 4 and 5 which contain the cell and marginal means provides an indication of the direction of the effects. Clearly, women who bought non-phosphate detergents were evaluated as being higher on the first and second factors and lower on the third factor than women who normally bought regular detergents. This appears to make sense since we might label the first factor as being something of a social concern or leadership factor. Thus the results indicate that people who purchase non-phosphate detergents are significantly higher in social concern or leadership than those who purchase regular detergents. Similarly, we might call the second factor a success factor where success is viewed in terms of goal attainment. The results would then indicate that people see users of non-phosphate detergents as being more successful than those who purchase regular detergents. Finally, the last factor which is an uninformed impulsiveness factor indicates that people who purchase non-phosphate detergents are seen as being more informed and less impulsive than those who buy regular detergents.

The absence of an interaction effect between the "purchase" factor and the "woman evaluated" factor has interesting implications. If a person's purchase and use of a product would influence his perception of users of that product along the dimensions relevant to that perception we would expect a significant symmetric interaction effect. In other words, we would expect those who buy regular detergent to rate other women who buy regular detergent higher on the social leadership and success factors and lower on the uninformed impulsiveness factor than women who buy non-phosphate detergents since the former two traits are presumably good and the last is probably seen as bad. Likewise we would expect the reverse to occur for those people who regularly purchase non-phosphate detergents. That this didn't occur provides evidence to indicate that the characteristics people attribute to others are stable percepts held in general by the population from which the sample was selected.

The lack of an interaction also provides some inferential evidence indicating that the characteristics we attribute to users of these two product types are veridical. If they were not we would not expect the characteristics attributed to users of a given product type by people who do not purchase it to be reinforced by people who do purchase it.

Summary

A study was conducted in which people were asked to attribute characteristics to users of regular detergent or users of non-phosphate detergent. It was found that significantly different characteristics were attributed to users of non-phosphate detergents than to users of regular detergents. The attribution of these traits was found not to be influenced by whether or not the respondent purchases the product in question.

References

- Calder, B. J. & Burnkrant, R. E. Attribution Theory and Consumer Behavior. In Final Preparation. Urbana-Champaign, Illinois.
Champaign-Urbana City Directory. Loveland, Colorado: The Johnson Publishing Company, 1972.

- Jones, E. E. & Davis, K. E. From Acts to Dispositions. In L. Berkowitz' (ed.), Advances in Experimental Social Psychology, Volume 2. New York: Academic Press, 1965.
- Jones, E. E. & Harris, V. A. The Attribution of Attitudes. Journal of Experimental Social Psychology, 1967, 3, pp. 1-24.
- Kelley, Harold H. Attribution Theory in Social Psychology. In David Levine (ed.), Nebraska Symposium on Motivation. Lincoln, Nebraska: University of Nebraska Press, 1967, pp. 192-238.
- Settle, Robert B. Attribution Theory and Acceptance of Information. Journal of Marketing Research, 1972, 9, pp. 85-88.
- Settle, Robert B., Faricy, J. H. & Warren, G. T. Consumer Information Processing: Attributing Effects to Causes. Proceedings: 2nd Annual Conference. Association for Consumer Research, 1971.

AN INVESTIGATION OF SOME ALTERNATIVES TO THE LINEAR ATTITUDE MODEL

Bobby J. Calder and Richard J. Lutz
 University of Illinois
 at Urbana-Champaign

The concept of attitude has been of central interest to both researchers and practitioners interested in consumer behavior. Investigations of attitudes offer the possibility both of understanding the cognitive processes of consumers and of predicting their future behavior. Recently, many researchers in this area have been especially attracted to a class of attitude models known as the expectancy-value approach. In contrast to the more traditional cognitive-affective-conative approach to the organization of attitudes, these models employ two elements, the likelihood (expectancy) of each belief making up an attitude and the worth (value or affect) associated with each belief (see Calder and Ross, 1972). If products are conceived of as bundles of attributes about which consumers may have a variety of beliefs, the expectancy value approach seems well suited to the analysis of consumer behavior.

Perhaps the most influential expectancy-value model has been proposed by Fishbein (1963, 1965, 1967). His model is given by the equation

$$A = \sum_{i=1} B_i a_i \quad [1]$$

where A is an attitude toward an object or action, B is the strength of belief i about the object or action, and a is the evaluative aspect of belief i. These components are usually assessed by semantic differential rating scales (Fishbein and Raven, 1962). Notice that this model postulates that beliefs are processed as the summation of the multiplicative combination of belief strength and affect. A similar expectancy-value model of this form has been presented by Rosenberg (1956, 1960) in which the expectancy component is treated as the "perceived instrumentality" of a value and the value component is represented by the "importance" of the value. These models have received the most empirical attention in the consumer behavior area (e.g., Bass and Talarzyk, 1972; Cohen, Fishbein, and Ahtola, 1972; Sheth and Talarzyk, 1972).

Other models within the same expectancy-value approach are certainly possible, however, Anderson (1971), for instance, has argued for the extension of his information integration model to attitude change. This model is written as

$$R = \sum_{i=0} w_i s_i \quad [2]$$

where R may be considered the attitude response, s is the scale value of source-communication i (which could be a belief) on the attitude dimension, and w is a weight reflecting the importance of i ($w_i s_i$ represents an initial attitude). Though clearly an expectancy-value approach, this model calls for averaging rather than summation ($\sum w = 1$).

Indeed, one problem with the expectancy-value approach is that there exists a large number of possible composition rules for combining the two

components, of which summation and averaging are just two instances. Consequently, in order not to be arbitrary, an expectancy-value model must be directly based on theoretical considerations regarding the cognitive structure from which the composition or cognitive processing rule may be derived. But this is precisely where current expectancy-value models are weakest. Fishbein offers an account of his model in terms of classical conditioning and generalization; Rosenberg relies simply on a general functionalist point of view; and Anderson employs only mathematical considerations. None of these rationales are sufficient to justify a particular form of the expectancy-value model. The purpose of this paper is to describe a new expectancy-value model developed by Calder and Lutz (1972) which does have an explicit basis in a representation of cognitive structure.

The Vector Model

Cognitive structure is represented in this approach as a two-dimensional metric space. One dimension represents an affective component (liking or favorableness) and the other represents a cognitive component (likely or probable). Any belief an individual possesses about a product is characterized by a value on each of these dimensions, i.e., as a set of coordinates in the cognitive space.

Our model is stated in terms of vector forms. The points representing beliefs in the cognitive space may be considered as vectors from the origin of the cognitive space. In terms of a model of cognitive process, the question now concerns how these belief vectors combine to determine attitude. An obvious structural solution is to think of attitude as a resultant vector whose coordinates are the sum of the affective coordinates ($\sum a$) and the sum of the cognitive components ($\sum b$) of the beliefs.

Attitude is thus conceived of as the resolution of the forces created by several specific beliefs. The magnitude of the attitude is simply the length of the attitude vector from the origin. Since we know the coordinates involved, it follows from elementary plane geometry that attitude is given by

$$A = [(\sum_{i=1} a_i)^2 + (\sum_{i=1} b_i)^2]^{1/2} \quad [3]$$

where A is analogous to the traditional attitude score, a is the affect associated with belief i, and b is the strength of the belief i. Unlike previous expectancy-value models, the vector model follows directly from our assumptions about cognitive structure.

Preliminary Research

Two studies (Calder and Lutz, 1972) have been conducted to compare the vector model with direct reports of attitude and with predictions yielded by the Fishbein and Anderson models. Both studies obtained standard semantic differential measures of attitudes toward buying non-phosphate detergents and of the affect and belief strength aspects of several beliefs concerning attributes of non-phosphate detergent. The first study used a sample of ninety-two University of Illinois business students. (For this sample a measure of the importance of each belief was also obtained in order to test

Anderson's model.) The results indicated that the vector model correlated .60 with reported attitude, and the Fishbein and Anderson models correlated .60 and .64 respectively.

The second study employed a convenience sample of forty-six women employed in University of Illinois offices. The correlations were, somewhat inexplicably, lower for this sample but still comparable to each other. The vector model correlated .33 with reported attitude and the Fishbein model .37.

In general, these results indicate that the predictive validity of the vector model is as strong as that of other expectancy-value models. In fact, it would seem that predictive validity is not an especially satisfactory criteria for evaluating such models. The class of possible composition rules which would sometimes produce high correlations is no doubt extremely large. This is, however, precisely why we believe the vector model and other cognitive structure approaches deserve serious attention. By having a basis in cognitive theory, they may be examined beyond their predictive ability.

Conclusions

In order to suggest some possibilities for research on the vector model, perhaps we should briefly note two other properties of the model. First, the model can be given a much more general form. Consider the Minkowski r -metric for determining distance in a metric space

$$d_{ij} = \left[\sum_{k=1}^n |x_{ik} - x_{jk}|^r \right]^{1/r}, \quad r \geq 1,$$

where each difference in coordinates on dimension k is raised to the r^{th} power, these are summed, and finally the r^{th} root is taken. Clearly the vector model, Equation [3], is a special case of this distance function with $r = 2$. It is thus possible to state a "city-block" ($r = 1$) version of the vector model. Therefore, r would seem to be an interesting parameter for further research. Likewise, the way in which the components (Σa) and (Σb) are weighted is also of interest. It can be shown (Calder and Lutz, 1972) that these weights may be written as $(\Sigma a/A)^{r-1}$ and $(\Sigma b/A)^{r-1}$ respectively.

Another feature of the vector model is that attitudes may be described not only by the property of magnitude but also direction. Some attitudes may be predominantly cognitive oriented whereas others may be more affective. This property suggests a new approach to the old problem of types of attitudes.

To summarize, traditional expectancy-value models are severely limited by the absence of any cognitive theory from which they may be directly derived. The vector model described here does possess a basis in cognitive structure. It appears to have predictive validity comparable to the other models. Moreover, the vector model seems to have several properties worthy of future research.

References

- Anderson, N. H. Integration theory and attitude change. Psychological Review, 1971, 78, 171-206.
- Bass, F. M. and Talarzyk, W. W. An attitude model for the study of brand preferences. Journal of Marketing Research, 1972, 9, 93-96.

- Calder, B. J. and Lutz, R. J. Cognitive structure and attitude models, manuscript, 1972.
- Calder, B. J. and Ross, M. Attitudes and Behavior. New York: General Learning Press, in press, 1972.
- Cohen, J., Fishbein, M., Ahtola, O. T. The nature and uses of expectancy-value models in consumer attitude research. Journal of Marketing Research, 1972, 9, in press.
- Fishbein, M. An investigation of the relationships between beliefs about an object and the attitude toward that object. Human Relations, 1963, 16, 233-240.
- Fishbein, M. A consideration of beliefs, attitudes, and their relationship. In I. D. Steiner and M. Fishbein (eds.), Current Studies in Social Psychology. New York: Holt, Rinehart and Winston, 1965, pp. 107-120.
- Fishbein, M. A behavior theory approach to the relations between beliefs about an object and the attitude toward the object. In M. Fishbein (ed.), Readings in Attitude Theory and Measurement. New York: Wiley, 1967, pp. 389-400.
- Rosenberg, M. J. Cognitive structure and attitudinal affect. Journal of Abnormal and Social Psychology, 1956, 53, pp. 367-372.
- Rosenberg, M. A structural theory of attitude dynamics. Public Opinion Quarterly, 1960, 24, pp. 319-340.
- Sheth, J. N. and Talarzyk, W. W. Perceived instrumentality and value importance as determinants of attitudes. Journal of Marketing Research, 1972, 9, pp. 6-9.

A CROSS-CULTURAL ANALYSIS OF HUSBAND-WIFE ROLES
IN HOUSE PURCHASE DECISIONS¹

Donald J. Hempel²
University of Connecticut

The mass media of our society reflect a growing interest in women's liberation, equal employment opportunities, and marital roles. In order to better understand the effects of these environmental changes on marketing opportunities, there is need for more comprehensive knowledge of male and female roles in household buying behavior. Marketing literature on family decision processes should be extended to provide a more reliable basis for predicting variations in husband-wife influence across product-market situations and to include comparative studies of international differences.

Knowledge of family role structures is of fundamental importance to the development of marketing theory and practice. The interaction of family members and their influence upon consumption decision processes are two of the basic determinants of buyer behavior. Family role differentiation has many implications for market segmentation and other strategy decisions, ranging from the selection of survey respondents to the formulation of communication programs. Despite the widespread recognition of family influence in many purchases, the decision-making processes and influence patterns within the household have received very little attention from marketers (Robertson, 1971; Frank, Massy, and Wind, 1972).

This paper explores several aspects of husband-wife interaction in the decision processes of recent house buyers in Connecticut and Northwest England. The analysis centers upon cross-cultural comparisons of the marital roles which were perceived by husbands and wives in five specific house-buying decisions. It also examines the consensus of role perceptions between the sexes and considers some of the potential determinants of family role structure in the purchase of a home.

Relevant Literature and Concepts

Research studies of family role structures by sociologists are abundant, but there is little consensus in the findings and conclusions reported. In a widely quoted survey of literature published during the last decade, Safilios-Rothschild observes:

Despite the availability of studies, one would have an impossible task, should he want to describe the power structure in the American family, unless he was willing to settle for just one major study. Actually, many family sociologists have done just this, mainly by relying on the findings of Blood and Wolfe's Detroit study (Blood and Wolfe; 1960). Not only have they settled for one study but also for one family member's perceptions of power structure, the wife's, and they perpetuate a "wives family sociology." [Safilios-Rothschild, 1970, p. 539]

She goes on to explain that the reasons for the lack of consensus are mainly methodological issues arising from the incomparability of decisions, the failure to include data from both spouses, and dependence upon global decision-making scores. Synthesis of research findings is further impaired by diversity in the conceptualization of power structure and by the controversy between advocates of survey and observational methods.

The evidence available from marketing literature is more comparable in decision focus, but much less abundant and similarly limited by conceptualization and methodological problems. Recent studies of family purchase decisions have presented summaries of these issues (Granbois, 1971; Davis, 1970) and have raised important questions about the reliability and validity of purchase influence measures (Davis, 1971). Although the empirical results of these recent studies are limited because of small convenience samples and significant separation from the actual decision process, they contribute both evidence and logic which challenge existing belief systems concerning family influences on consumer behavior.

Several recurring arguments regarding conceptual approach and method can be found in the current literature of both marketing and sociology. First, family power or role structure should be treated as a multidimensional concept which can be measured both through the outcomes and through the process of decision-making. This requires a sharper distinction between the concepts of power, influence, authority, and responsibility. Measures of authority and responsibility can be operationalized through the outcome-oriented question, "Who decides?" An operational measure of influence is more difficult because it is process-oriented and must reflect subtle pressures within the family. Second, decision-making in the family, particularly for major economic purchases, should be viewed as a multiphase process involving a number of specific and related decisions. This approach requires measures of influence or responsibility for each decision in order to identify role differentiation by decision type. It also discourages reliance upon an overall measure of family role structure, such as the Blood and Wolfe index. Third, family decision-making should be approached as a small group process involving the husband-wife dyad and perhaps children or other household members. Since the extent of involvement and the role perceptions of each family member vary across households and decisions, multiple-person measurements within each family are needed. This approach requires separate measures of the decision roles perceived by the main participants, particularly both spouses.

The nature of husband-wife interaction in a buying process can be expressed in terms of whether a specific decision is shared or dominated by one spouse. This concept of dominance has been widely used by researchers in marketing and sociology, but the meaning attached to the term is not always explicit and is often varied (Komarovsky, 1958; Blood and Wolfe, 1960; Heer, 1963; Davis, 1970; and Granbois, 1971). In most of these studies the concept has been used to describe the outcomes of the family decision process - that is, who finally makes a particular decision. Two recent studies have attempted to extend the concept of dominance (Kelly and Egan, 1969) and validate alternative measures of family power structure (Olson and Rabunsky, 1972). These investigators encourage the study of both the decision outcome and the family process through which the decision is reached. If family decision-making is to be treated as a joint reciprocal activity involving discussions and negotiation, then a more dynamic conceptual model, dealing with both process and outcome, should be utilized. This approach is very appropriate for studies of major purchases, but it is difficult to implement

through survey research methods. Some combination of survey and observational techniques is probably necessary to best measure both aspects of family decisions, particularly if a multidimensional concept of role structure is to be used.

This paper attempts to incorporate some of these conceptual and methodological considerations within a cross-cultural study of the decision processes involved in house buying. The analysis centers upon the husband-wife role differentiation for selected decisions within this purchase process. It is believed that the purchase of a home involves a decision-making process which is likely to reveal the most basic role structure of the family as it relates to consumption expenditures. This product class stimulates high personal involvement because of the infrequency of purchase and its very important social and economic consequences for all household members. The complexity of the decision process encourages role specialization, but the active participation and influence of both spouses is fostered by the importance of the decision. Given the differences in values, knowledge, and preferences which each spouse brings to focus on the decision, some conflict inevitably results. One would expect a preponderance of joint decision-making to result from the high product saliency. However, the decision complexity is likely to mitigate this pattern and produce task segregation and perceptual differences. Some prior evidence supporting these expectations is provided by a small exploratory study of house buying in the Vancouver, B.C. metropolitan area (Kelly and Egan, 1969).

Methodology

The data for this study were obtained from two comprehensive surveys of recent house buyers conducted during the summers of 1968 and 1971. These studies were designed to yield comparable measures of consumer decision processes in different housing markets. The first sample involved 206 families from Hartford, Connecticut and eight adjacent towns. The 1971 study included 317 households from seven towns in the Preston-Lancaster area of Northwest England. Both investigations incorporated probability samples of households who purchased either a new or previously occupied house and recorded their ownership during the first six months of the study year.

The data collection procedures involved extensive personal interviews in each household, followed by a request that two self-administered questionnaires be completed separately by the husband and by the wife and returned by mail. The importance attributed to the subject by the respondents was reflected in a response rate of 77 percent for the mail questionnaires in Connecticut and 67 percent in England. The data collection instruments used were identical in both studies, except for some minor adaptations to English word usage in the 1971 study. Details concerning the research design and questionnaires used in Connecticut have been published elsewhere (Hempel, 1970).

The concept of dominance used in this study was operationally defined by asking a series of questions about the relative importance of each spouse in selected house-buying decisions. Two different measures of husband-wife roles in the decisions process were contained in the separate self-administered questionnaires distributed at the time of the interview. The husband's questionnaire asked him to indicate who was the decision-maker "mainly responsible" for the set of decisions shown in Table 1. The questions used to measure the wife's role perceptions asked her to indicate the "relative influence" which she and her spouse had "during the process of buying" in each of the decisions shown in Table 2. Thus, the wife's response provides a measure of perceived

Table 1

Marital Roles in Selected House Buying Decisions
as Perceived by Husbands

Decisions	Relative Responsibility Perceived by Husband (percent of respondents ^a)					
	Hus. Dominant		Joint Decision		Wife Dominant	
	Conn	NWEng	Conn	NWEng	Conn	NWEng
1. Neighborhood	6	12	79	74	16	14
2. Style of house	7	12	67	71	26	17
3. When to purchase	29	28	61	67	10	6
4. Acceptable price	39	35	59	61	2	4
5. Where to apply for mortgage	62	50	34	45	4	5
6. Who to contact for assistance in finding house	26	33	55	58	19	9
7. Who to rely on for advice in financing purchase	53	43	43	54	4	3

^aN₁ = 154 for Connecticut and N₂ = 195 for Northwest England

decision influence in the decision process, whereas the husband's answers indicate his perception of decision-making responsibility or identification of the primary decision maker. These are conceptualized as measures of two closely related but different dimensions of family role structure.

The responses of both husband and wife were scored by assigning a value of "1" when the wife's role was dominant, "2" for equal or joint decisions, and "3" for answers indicating that the husband's role was dominant. On the basis of these individual decision scores on a 3-point scale, an overall measure of perceived dominance was computed for each spouse. This aggregate measure included only the first five decisions which were common to both questionnaires (when, price, style, neighborhood, and mortgage), so its value could range from 5 to 15. Respondent attribution of greater importance to the husband's role across the five decision areas is reflected in higher values for the dominance index. This overall decision-making score provides a single measure of the husband's relative importance in the purchase decision. It is a useful means of supplementing the individual decision data with a summary value to simplify analysis of possible role determinants. However, it would be misleading to rely solely upon the analysis of trends in this index, because the husband-wife influence patterns do differ by type of decision.

Findings

The analysis and results of this study concerning husband-wife influence are presented below in three parts. First, the differences in roles perceived by the marital partners are examined across a set of major decisions within the home-buying process. Next, the extent of agreement between the role perceptions of husband and wife is considered for the five decision areas evaluated by both respondents. Finally, some of the possible determinants of these decision patterns are analyzed as bases for market segmentation.

Role Differentiation with the Buying Process

Table 1 shows the distribution of husbands' responses to questions about the relative decision responsibility of each spouse for seven house-buying decisions. Table 2 shows the distribution of wives' responses to questions about the relative influence of each partner in a similar set of decision areas. The first five decisions in each table are listed in order of decreasing proportions of joint decisions based on the average decision scores of all Connecticut respondents. The last two decisions in each table are segregated from this ordering because data were obtained only from the husband or the wife for these questions.

Do the marital roles differ with regard to the type of decision? There are substantial differences in the relative influence of husbands and wives across the decision areas examined. Data from both Connecticut and England indicate that the dominance of the husband was least likely to occur in decisions concerning neighborhood and house style, and most likely in financial decisions involving price and mortgage source. These data provide some support for the hypothesis of role specialization, suggesting that the husbands concentrate on instrumental-financial decisions while the wives give more emphasis to expressive-social dimensions. However, the modal pattern for all decisions except those involving mortgage financing indicates that house-buying is a shared decision process. The percentage of joint decisions ranged from 90 percent in the choice of house size (English wives) to 34 percent in the decision about mortgage source (Connecticut husbands).

Table 2

Marital Roles in Selected House Buying Decisions
as Perceived by Wives

Decisions	Relative Influence Perceived by Wife (percent of respondents ^a)					
	Hus. Dominant		Joint Decision		Wife Dominant	
	Conn	NWEng	Conn	NWEng	Conn	NWEng
1. Neighborhood	9	6	78	80	13	14
2. Style of house	6	4	77	84	17	12
3. When to purchase	25	16	59	80	15	4
4. Acceptable price	36	21	57	74	7	5
5. Where to apply for mortgage	47	40	45	57	8	3
6. Size of house	6	5	79	90	15	5
7. Floor plan and layout	5	7	77	79	18	14

^aN₁ = 152 for Connecticut and N₂ = 196 for Northwest England

Do the marital roles in house-buying differ with regard to the cultural setting of the respondents? There is considerable cross-cultural consistency indicated by the rank-order of the decision roles perceived by both sexes. For the five major decisions, there was perfect agreement in the ordering of decisions according to the proportion of families in which the husband's role was perceived to be dominant. When the distribution of the husbands' response scores among the three role categories was compared for each cultural group, none of the international differences in their perceptions were statistically significant at the .05 level. However, a different interpretation is indicated by cross-cultural comparisons within a particular role category. Connecticut husbands reported more frequently than their English counterparts that they were mainly responsible for decisions about mortgage source (62 percent vs. 50 percent) and sources of financial advice (53 percent vs. 43 percent). They were somewhat less likely than English males to claim dominance in decisions about house style and neighborhood and were more inclined to attribute responsibility for these areas to their wives.

There was less intercultural similarity in the distribution of roles reported by wives, although women in both countries did indicate a similar rank-ordering of decisions in which male influence dominated. A Chi-Square test of cultural similarity in the wives' response distributions revealed that the role structures perceived by English and American women were significantly different for decisions about when to buy ($\chi^2 = 21.80$, $p < .001$), price ($\chi^2 = 10.70$, $p < .01$), mortgage source ($\chi^2 = 7.16$, $p < .05$), and size of house ($\chi^2 = 10.37$, $p < .01$). Both sexes perceived more sharing of decision roles in England than in Connecticut. Cross-cultural differences in the proportion of families reporting joint decisions ranged from several percentage points for decisions about neighborhood and style, to a highly significant discrepancy of 17 percentage points for the price decision (wives' responses).

Consensus of Role Perceptions

Do husbands and wives, considered as groups, differ in their perceptions of decision roles? As noted earlier, the comparisons of husband and wife responses in this study are conceptualized as involving two dimensions of role structure--decision responsibility and influence. Agreement between the role perceptions probably reflects a higher level of marital consensus than a single measure of role structure, such as "who decides?" The level of agreement was ascertained by comparing the husbands' responses in Table 1 with those of the wives' in Table 2 for the first five decision areas. These comparisons could be simplified by presenting a table of mean scores, but this aggregation would tend to conceal differences in the distribution of responses across the role categories.

The comparisons of husbands and wives as groups indicate that there is very little consensus of role perceptions within either country. A Chi-Square test was made for each of the 3-by-3 contingency tables containing the distributions of husband and wife responses on the 3-point scales used to measure role perceptions. These tests revealed significant sexual differences among the Connecticut respondents for decisions about style of house ($\chi^2 = 18.81$, $p < .001$), when to purchase ($\chi^2 = 38.14$, $p < .001$), and mortgage source ($\chi^2 = 42.94$, $p < .001$). The differences in role perceptions of husbands and wives in the English sample were significant for all of the decisions considered: neighborhood ($\chi^2 = 42.20$, $p < .001$), style of house ($\chi^2 = 12.24$, $p < .05$), when to purchase ($\chi^2 = 10.11$, $p < .05$), price ($\chi^2 = 40.31$, $p < .001$), and mortgage source ($\chi^2 = 38.16$, $p < .001$).

Wives in both countries were more likely than husbands to perceive joint decisions about the style of house and mortgage source. The majority of husbands reported that they were responsible for the mortgage decision, and there was a tendency for them to attribute the house style decision to their wives. Disagreement in the perceived role structures among English families is particularly evident in the percentage of respondents who reported joint decisions. The proportion of shared decisions reported by wives exceeded that for husbands by an average of 11 percent in England versus 5 percent in Connecticut. If sexual differences in mean decision scores were to be used as a basis for this analysis, the level of agreement would appear to be higher in England than in Connecticut. Comparisons of mean scores tend to provide misleading results, because the response distributions for English wives were so heavily concentrated in the joint decision category. The extent of husband-wife agreement in role perceptions can be examined more effectively on the basis of intrafamily comparisons.

Do husbands and wives within the same family differ in their perceptions of role performance? The data contained in Table 3 indicate the extent of role consensus within the individual family, determined by the differences in the 3-point decision scores computed for each spouse. A significant disagreement in perceptions is indicated by the difference of one scale point that results when a joint decision is perceived by only one partner. A difference of two scale points reflects complete polarity of perceptions, with one spouse reporting a husband-dominant decision while the other indicates that it is wife-dominant.

The intrafamily comparisons indicate that the consensus of role perceptions ranged from 76 percent for the English neighborhood decision to 57 percent for the Connecticut price decision. Marked similarity in the percentage of couples who agree about their roles can be seen in the data for both countries, particularly in the decisions about house style, when to purchase, and mortgage source. Connecticut couples were less likely than English respondents to agree about the price (57 percent vs. 68 percent) and neighborhood (70 percent vs. 76 percent) decisions. Chi-Square analysis indicated that the cultural differences in the distribution of discrepancy scores were not significant (at the .05 level) for any one of the five decision areas considered in Table 3.

Disagreements in role perceptions were confined almost entirely to situations in which one spouse perceived a joint decision while the other attributed more importance to one partner. Polarity in role performance was reported by not more than 4 percent of the households in either country. The shape of the discrepancy score distributions tends to be symmetrical, but there is some indication that the disagreements favor the spouse who is generally considered to be the less dominant decision influence. For example, there is a tendency for the husband to attribute more importance to the wife than she attributes to herself in decisions about mortgage, price, and when to purchase, and to exhibit a self-bias for decisions about style and neighborhood. Of course, this same pattern would emerge if the wife were understating her role in the financial-scheduling decisions and overstating her importance in the style and neighborhood decisions. In either case, these compensating role allocations may reflect reporting errors which result from the respondent's efforts to be conventional or adhere to socially desirable norms. Such behavior patterns could occur if the husband were motivated by equalitarian considerations in his reporting of decision responsibility, while the wife attempted to present an influence pattern which was consistent with cultural expectations of role specialization.

Table 3

Consensus of Marital Role Perceptions in Selected House-Buying Decisions

Decisions	Extent of Husband-Wife Agreement in Role Performance (percent of respondents ^a)									
	Husband attributes <u>less</u> importance to wife than she attributes to herself.		Husband and wife agree.		Husband attributes <u>more</u> importance to wife than she attributes to herself.					
	Conn	NWEng	Conn	NWEng	Conn	NWEng	Conn	NWEng		
	-2	-1	0	+1	+2					
1. Neighborhood	1	17	9	70	76	10	12	1	2	
2. Style of house	1	0	18	17	68	68	11	13	1	2
3. When to purchase	b	b	12	15	66	64	19	18	3	2
4. Acceptable price	b	b	17	10	57	68	24	21	b	2
5. Where to apply for mortgage	0	0	8	12	67	67	24	19	1	1

^aN₁ = 145 for Connecticut and N₂ = 185 for Northwest England ^bLess than one percent

Some Possible Determinants of Role Structure

What are the bases or determinants of decision-making role structures? The major school of thought concerning family influence explains marital role structures in terms of the relative resources contributed to the household by each individual. This "theory of resources" recognizes contributions in the form of both conventional socio-economic dimensions (income, education, and occupation) and more subtle socio-psychological inputs (personal involvement, affection, and performance of household tasks). The explanation of why one spouse is dominant rather than the other has been troubled by the lack of authority and influence data and their interrelations with whoever makes the decision (Safilios-Rothschild, 1970).

This section of the paper examines an extensive set of potential explanatory variables in terms of their relation to the dominance index computed for each spouse. Correlation and regression analysis are used to explore these relationships without rigorous adherence to the assumptions underlying the techniques. Many of the variables (including the dependent measures of dominance) are not intervally scaled, and there are nonlinearities and interactions which could be handled through transformations or dummy variables. Given the inconsistency of available evidence and the absence of theoretical guidelines in this field of study, such adjustments would be largely arbitrary. The results of this analysis should be interpreted cautiously and considered as exploratory evidence useful in the construction of a more refined model.

The correlation coefficients presented in the first four columns of Table 4 indicate that there is considerable cross-cultural consistency in the relationships examined. In general, the degree of association between the explanatory variables and the dominance indices was somewhat higher in Connecticut than in England and greater among husbands than among wives. The age of both spouses and the number of years married were inversely related to the relative importance of the husband, as perceived by both partners. Marital independence, represented in the proportion of years one was single and an early life cycle stage, were positively related to male dominance in Connecticut and England respectively, but the degree of relationship was less consistent across cultures.

The education of both husband and wife, and the husband's occupational status were positively correlated with all of the dominance indices, while the wife's educational status was negatively correlated. The relation of educational and occupational differences between spouses (as measured by status consistency ratios) to the dominance indices indicates that the importance of the husband's role was directly associated with the extent of his relative status advantage, especially in Connecticut. The curvilinear hypothesis that women tend to be more dominant in the lower classes while men are more dominant in the upper classes is supported by the data from both countries; this pattern is most evident in the wife's perceptions of role structure. Family income was positively associated with greater male dominance, but the degree of relationship was relatively low for Connecticut husbands.

Measures of the wife's employment status indicate that the perceived importance of the husband's role declined as his spouse's commitment to work outside the home increased. The wife's contribution to family income, the number of hours she was employed, and the years she had worked with the firm were all negatively correlated with the dominance indices. These changes in the wife's employment status were highly significant in Connecticut, but they appeared to be of less importance in England. The cultural differences were more evident in the role perceptions of the husband.

Table 4

Potential Explanatory Variables for a Model of Family Role Structure

Variable Name and Description	Correlations with Dominance Index				Stepwise Analysis of Dominance Index-- Partial Regression Coeff. and (t-Ratios)			
	Husbands		Wives		Husbands		Wives	
	Conn	NWEng	Conn	NWEng	Conn	NWEng	Conn	NWEng
Husband's age (years)	-.135	-.209	-.067	-.137	.048 (2.31)	-.035 (2.75)	.050 (2.11)	
Wife's age (years)	-.168	-.176	-.106	-.140				
Age consistency (husband's age + wife's age)	.094	-.028	.102	-.007				
Years married	-.259	-.153	-.190	-.059	-.074 (2.92)		-.099 (3.38)	
Husband's independence index (age + years married)	.228	.050	.197	.088				
Wife's independence index (age + years married)	.226	.036	.201	.072				
Husband's education (years)	.224	.147	.312	.169	.089 (1.89)			-.077 (1.12)
Wife's education (years)	.111	.154	.078	.093			.100 (1.58)	
Educational consistency (husband's education + wife's education)	.119	-.008	.241	.123			2.407 (3.45)	
Husband's occupational status (scaled 3-12)	.157	.097	.249	.314	-.363 (3.24)			
Wife's occupational status (3-12)	-.284	-.086	-.076	-.056	.213 (1.97)			
Occupational consistency (husband's occupation + wife's occupation)	.351	.129	.222	.319	1.679 (3.41)			.369 (2.85)
Life cycle I - Young, no dependents (0,1)	.004	.135	.033	.180	.725 (1.96)		.837 (2.67)	
Life cycle II - Young, dependents (0,1)	.096	.058	-.008	-.067				
Social class I - high educational and occupational status (0,1)	.125	.123	.195	.239				.621 (1.59)
Social class II - low educational and occupational status (0,1)	-.192	-.085	-.281	-.301				
Size of household (number of persons)	.057	-.035	.050	.014				.141 (1.57)
Annual family income (\$ or £)	.047	.144	.230	.179		.437 (2.30)	.093 (3.05)	
Wife's contribution to income (%)	-.235	-.033	-.105	.036	-.011 (1.33)		-.015 (1.58)	
Hours wife employed outside home	-.290	-.056	-.056	.066		.010 (1.17)	.012 (1.22)	
Years wife employed by firm	-.273	-.118	-.165	-.059				
Husband's involvement index (sum of 15 attitude scores)	-.090	.189	-.124	-.157			-.011 (1.23)	-.011 (1.30)
Wife's involvement index (sum of 15 attitude scores)	-.037	.184	.012	-.065		.032 (2.41)		
<u>Regression Analysis Data</u>								
Intercept					7.61	9.63	5.31	10.46
Degrees of freedom					128	85	122	74
Standard error of estimate					1.15	1.48	1.32	0.93
R ²					0.284	0.146	0.248	0.219

The intensity of personal involvement in the house-buying decisions was measured by an attitude index. Separate involvement indices were constructed for each spouse by aggregating their scaled responses to a series of 15 questions dealing with the relative importance of various reasons for buying a home. High involvement of either spouse was positively associated (in England) with the husband's perception of role structure, but negatively correlated with the importance of the husband as perceived by the wife.

The regression results shown in Table 4 were developed through a step-wise regression technique which eliminated independent variables from the final equation when the t-ratios for their partial regression coefficients were not 1.0 or greater. This procedure was used to screen out relatively unstable and less efficient predictors of the dependent measures from the set of 23 potential explanatory variables. These data provide a preliminary estimate of the structural relationships among a reduced set of explanatory variables and the scaled dominance indices.

The regression analysis indicates that basic demographic measures of age, years married, education, occupation, and wife's employment status are useful predictors of perceived role structure. The inclusion of less conventional measures which represent educational and occupational consistency, and the intensity of personal involvement appears to add measurably to the predictive power of the equation. The R^2 values in the range of 15 percent to 28 percent are encouraging, because the predictions of the model can be improved through the introduction of transformations and other independent variables. For example, detailed analysis of the wife's employment status variables has revealed considerable interaction between the effects of these measures and those of family income. The presence of a working wife is associated with the perception of relatively low husband dominance by wives in the lower income groups, a sharp increase in the mean dominance index for the middle income families, and a fairly small increase (Connecticut) or decline (England) for the upper income households. The perceptions of role structure among husbands exhibit a similar pattern in Connecticut, but the wife's employment outside the home is associated with a very high level of husband dominance among the lower income groups in England. Most of the work in process on this study involves the construction of indices and composite variables to represent such interaction patterns more effectively in refined multivariate models.

Summary and Conclusions

This empirical study of family role structure provides evidence that decision task specialization and perceptual differences do exist among some recent house buyers in two different cultural settings. There was considerable intercultural similarity in the rank order of the husband's relative importance across a set of major house purchase decisions. Joint decision patterns predominated in both countries, but dominance by one spouse was perceived more frequently by Connecticut buyers. Differences in role perceptions between sexes within each country were generally greater than the international differences. A number of basic household characteristics were consistently related to variations in perceived role performance among spouses in both countries.

The results of this study indicate clearly that perceptions of both husband and wife should be measured independently if one wishes to explain marital roles in major purchase decisions. The conceptual and methodological issues which plague family influence studies should be clarified by more integration of research in marketing and sociology. Improved conceptualization and operationalization of influence, authority, and power concepts are needed to develop

more sophisticated and valid theoretical structures in both disciplines. More recognition of perceptual differences within the family and of the specificity of decisions within the purchase process is essential to the development of buyer behavior theory.

Footnotes

1. This research was financed in part by grants from the University of Connecticut Research Foundation, the Marketing Department of the University of Lancaster (England), the Center for Real Estate and Urban Economic Studies at the University of Connecticut, and the National Science Foundation Grant GJ-9 to the University Computer Center.
2. Associate Professor of Marketing, University of Connecticut.

References

- Blood, Robert O. & Wolfe, Donald M. Husbands and Wives: The Dynamics of Married Living. Illinois: The Free Press, 1960.
- Brown, G. H. The Automobile Buying Decision Within the Family. In N. N. Foote (Ed.), Household Decision Making. New York: New York University Press, 1961.
- Davis, H. L. Measurement of Husband-Wife Influence in Consumer Purchase Decisions. Journal of Marketing Research, 1971, 8, 305-312.
- Davis, H. L. Dimensions of Marital Roles in Consumer Decision Making. Journal of Marketing Research, 1970, 7, 168-77.
- Davis, H. L. Some Problems in Building a Propositional Inventory: The Case of Family Decision Making. Paper presented at the Conference of the American Marketing Association, September 1968.
- Engel, J. F., Kollat, D. T. & Blackwell, R. O. Consumer Behavior. New York: Holt, Rinehart, 1968.
- Ferber, Robert. On the Reliability of Purchase Influence Studies. Journal of Marketing, 1955, 19, 225-32.
- Frank, Ronald E., Massy, William F. & Wind, Yoram. Market Segmentation. New Jersey: Prentice-Hall, 1972.
- Frey, J. N. & Siller, F. R. A Comparison of Housewife Decision Making in Two Social Classes. Journal of Marketing Research, 1970, 7, 33.
- Granbois, Donald H. & Willett, Ronald P. Equivalence of Family Role Measures Based on Husband and Wife Data. Journal of Marriage and the Family, 1970, 32, 68-72.
- Granbois, Donald H. A Multi-Level Approach to Family Role Structure Research. In David M. Gardner (Ed.), Proceedings of the Association for Consumer Research, 1971, 99-107.
- Heer, David M. The Measurement and Bases of Family Power: An Overview. Marriage and Family Living, 1963, 23, 133-9.
- Hempel, Donald J. A Comparative Study of the Home Buying Process in Two Connecticut Housing Markets. Connecticut: Center for Real Estate and Urban Economic Studies, The University of Connecticut, 1970.
- Jaffee, Laurence J. The Roles of Husbands and Wives in Purchasing Decisions. In Lee Adler and Irving Crespi (Eds.), Attitude Research at Sea. Chicago: American Marketing Association, 1966.
- Kenkel, William F. Husband-Wife Interaction in Decision Making and Decision Choices. In Martin M. Grossak (Ed.), Understanding Consumer Behavior. Boston: The Christopher Publishing House, 1964.

- Komarovsky, M. Class Differences in Family Decision Making on Expenditures. In N. N. Foote (Ed.), Household Decision Making. New York: New York University Press, 1961.
- Komarovsky, M. Family Buying Decisions: Who Makes Them, Who Influences Them? Printer's Ink, 1958, 22-28.
- Olson, David H. & Rabunsky, Carolyn. Validity of Four Measures of Family Power. Journal of Marriage and the Family, 1972, 34, 225-34.
- Robertson, Thomas S. Innovative Behavior and Communication. New York: Holt, Rinehart and Winston, 1971.
- Safilios-Rothschild, C. A Study of Family Power Structure: A Review of 1960-1969. Journal of Marriage and the Family, 1970, 32, 539-552.
- Safilios-Rothschild, C. The Influence of the Wife's Degree of Work Commitment Upon Some Aspects of Family Organization and Dynamics. Journal of Marriage and the Family, 1970, 32, 681-691.
- Safilios-Rothschild, C. Family Sociology or Wives Family Sociology? A Cross-Cultural Examination of Decision Making. Journal of Marriage and the Family, 1969, 31, 290-301.
- Sharp, Harry & Mott, Paul. Consumer Decisions in the Metropolitan Family. Journal of Marketing, 1956, 21, 149-56.
- Wolgast, Elizabeth H. Do Husbands or Wives Make the Purchasing Decisions? Journal of Marketing, 1958, 23, 151-8.
- Zober, Martin. Determinants of Husband-Wife Buying Roles. In S. Britt (Ed.), Consumer Behavior and the Behavioral Sciences. New York: Wiley, 1966.

CROSS-CULTURAL CONSIDERATIONS IN CONSUMER BEHAVIOR:
THE CASE OF THE CONSUMER IN THE SOVIET UNION

Thomas V. Greer
University of Maryland

The purpose of this paper is to sketch the Soviet consumer in a context largely ignored in the West and largely ignored in Soviet information aimed at the West. The environment typically presented is that of Moscow, Leningrad, and three or four other large cities in the European portion of the U.S.S.R. However, the author is interested here in the other Soviet consumers: the rural Europeans; the Asian minorities; and the frontier settlers. It is hoped that this material will put into perspective much that is already known and understood about the Soviet urbanite.

Although the author has made trips to the U.S.S.R. and conducted research there, and had valuable assistance from the Soviet Embassy in Washington, D.C., the methodology, to a considerable extent, consists of a careful analysis and synthesis of Russian-origin literature. Admittedly not ideal, this methodology is nevertheless standard practice among scholars interested in the Soviet Union in every academic discipline. The reasons for this are obvious.

The Rural European Soviet Union

According to the 1970 Census about 44 per cent of the Soviet population is rural, down from 52 per cent in 1959 and 68 per cent in 1939. However, the 44 per cent figure is misleading in that the definition of rural varies from republic to republic in the U.S.S.R. In addition, one must note that the population cutting points between the urban and rural categories in all republics are higher than in the United States. Furthermore some Soviet communities that are large enough to be classified urban are not so classified because they have a majority of agricultural workers. This is perhaps a good example of the potential danger in utilizing international data without extreme care. At any rate, the rural population remains large and significant. Reconciliation of the Soviet population data with those of the United States must wait for additional detail to be released by the Russians.

Russian farmers have always been a severe problem for the communist planners and leaders. It is not just the matter of productivity and meeting national needs for commodities. Just as important, if not more so, has been a philosophical difference in what the rural people aspired to and what the communist theorists felt the rural people should aspire to. What is more, communism was premised on the priority well-being of industrial workers, not farmers. Farmers were often considered petty bourgeoisie, along with merchants and individual artisans. Even in today's Soviet press, the workers and the peasants are normally referred to as "two different classes." Chairman Leonid Brezhnev, in his report to the 24th Communist Party Congress in March, 1971, referred to three classes in Soviet society: workers, peasants, and intelligentsia (Brezhnev, 1971). Rather than glorifying the agricultural sector, many government and party leaders and bureaucrats have tended to vilify it. Some westerners may recall that the Communist Manifesto of 1848 referred to the "idiocy of village life."

In Karl Marx's time and today the peasants are labeled reactionary in that they exhibit more individualism than do workers. (This is still less than what West European peasants show.) With a central thesis of large-scale production, Marx and his disciple Lenin had no basic sympathy for or empathy with the man of the soil or this characteristic. However, Lenin and his followers were pragmatic enough to appease the peasantry when setting up the bolshevik regime and for several years thereafter. Collectivization had to come much later. Joseph Stalin once told Winston Churchill that his struggle with the Russian peasants had been a more perilous and formidable undertaking than the battle for Stalingrad (Mitrany, 1961).

Paradoxically, Russian farmers traditionally exhibited some aspects that one is tempted to term early collectivism. Prior to 1917 many thousands of farms were owned by several peasant families jointly. Households generally performed their own work but sometimes shared work. But major decisions affecting the farm were jointly reached. This Russian institution influenced early Marxist thinkers to some extent, and had an effect on Mexico in the 1920's and 1930's, modern Israel over a long period of time, and India during the late 1940's and the 1950's.

The typical farm in the U.S.S.R. has 1,000 to 3,000 residents, and all able bodied persons above age 15 work. This work is physically hard and the economic results discouraging. Although growing, mechanization is spotty. The scythe and the ox are almost as common as the combine and the tractor. Living conditions and sanitation are primitive and much inferior to those of the city, while access to retailing facilities and services is extremely limited. Soviet economist A. Ilyin (1972) asserts that some of the most important problems in his country for the next few years are retention of young persons on the farm, improving the income of rural people, and providing skilled manpower in rural areas.

One can hardly think of the Soviet farmer without thinking also of the farmers' marketplace. Because of well known shortcomings of distribution channels, procurement, and general economic planning, this institution continues to flourish. There are 40 in Moscow alone. During the May to October period the farm is normally represented in marketplaces by one or more persons, often elderly and unable to perform hard labor. In addition, some other persons, also often elderly, are there selling their own products and perhaps those of several other households from the small private plots. Although the marketplace is also important as a social outlet, it is the economic that is overriding. Soviet economic historian N. Bromlei has concluded that "Profits from private plots constitute 30 to 50 per cent of the actual earnings of a collective farmer." (Bromlei, 1966)

The Asian Minorities

The precise extent to which the Asian minorities exercise an effective demand for distinctive or at least modified goods and services is by no means clear. Neither is it clear to what extent they constitute a different target for economic communications. Nevertheless, it is clear that the extent is noteworthy.

Soviet Islam embraces the major portion of the many non-European minorities. Thus, the generalizations one can make about Moslems in the U.S.S.R. go a long way toward characterizing the minorities. The approximately 40 million Moslems are found principally in Soviet Central Asia (Turkmenistan,

Uzbekistan, Tadzhikistan, Kirgizia), Kazakhstan, Bashkiria, the Tatar area along the Volga, and parts of the Caucasus. The consumption differences should be approached cautiously and must be estimated rather grossly and indirectly by studying cultural phenomena.

Nevertheless, Bennigsen and Lemerrier-Quelquejay (1966) state:

As a general rule . . . the Muslims, and above all those of Central Asia, appear to be more attached to their material way of life than they are to their religious traditions, and they only abandon it if it is incompatible with material progress. Most of the time, the Muslim does not resort to a Russian substitute unless it is more convenient, and even then he tries to effect a compromise between tradition and innovation (p. 200).

Such fundamental categories of goods and services as food, clothing, and shelter are revealing. Such things are slow to change. Cultural anthropologists have established that, in such categories, an innovation "must usually be immediately demonstrated as a superior substitute (Keesing, 1958, p. 411)." The dietary habits of the Moslems in the U.S.S.R. have shown strong resistance to massive change. They still consider pork unclean and seldom raise hogs on their collective farms. They often encounter pork in factory cafeterias and military mess halls, where there is seldom a substitute available. The great majority continue to avoid alcoholic beverages, although wine is gaining in favor, due in part to sizeable promotional efforts. The diet stresses heavy consumption of lamb, rice, unleavened bread, dairy products, and fresh and dried fruits. Pasta-type products are generally of logical origin rather than European. Nevertheless, there has been some very limited adoption of potatoes, European-type breads, and cabbage, an occurrence more common in urban than rural areas. The relative slowness with which dietary habits change is, of course, world-wide. It is all the more feasible in the Moslem regions of the Soviet Union because most of their food is grown and processed nearby. On the collective farms, too, Moslem dietary habits can be indulged fairly easily because such farms are usually all Moslem, except for one or two Russian technicians, or all European. In addition, the collective farmer has his private plot. The sovkhoses, or state farms, are somewhat less segregated.

The distinctiveness of clothing is showing a slight decline but is still extremely important. This is possible mainly because so much apparel is made at home or made to order by tailors. Textile factories turn out quite a lot of ethnically distinctive fabric, but apparel factories turn out relatively little ethnically distinctive merchandise. The traditional clothing of the minorities heavily predominates in the countryside and among women, whether urban or rural. The great majority of Moslem men strike a compromise, wearing some European clothing and some traditional and commonly mixing the two. There is extremely wide use of the knee-length outer robe, the ribbon sash, and countless variations of indigenous headgear. Urban teen-age boys demonstrate the greatest acceptance of European clothing, but even they often mix the styles. It is noteworthy that contemporary Soviet Central Asian art, especially painting, almost invariably depicts the subjects in traditional attire.

Shelter is a mixed set of phenomena. In the countryside newer housing normally uses the neat configuration of the Russians, one street or two streets crossing at right angles, rather than the seeming randomness of Moslem tradition in these regions. Wood has gained in popularity as sun-dried home-made brick has declined. There may be no floor. Within the house the beds are often only ornamental and the members of the family may sleep on the floor.

In both city and rural areas, it is common for one room to be furnished in a more or less European style for receiving European guests while the rest of the house or apartment is distinctively Asian. Of course, in the cities, many Moslems now live in apartment complexes, so that shelter differences are confined to the furnishings of the dwelling units. Nevertheless, some towns exhibit a "colonial" character. Although ethnic segregation of towns is not ubiquitous, it does occur. There may be a Moslem town and attached to it a Russian administrative town.

Consumption among Soviet Moslems is conditioned by several other factors of importance. Although there has been some decline, the family is still largely patriarchal. The wishes of the husband and father are paramount. They expect, and usually get, obedience from female members of the family. Equality of the sexes, taught so carefully in European districts of the U.S. S.R., has made little headway in these regions. The male members of the family tend to expect conservatism when they delegate authority to females or children. The practice of the father choosing his son's bride has begun to decrease only in recent years. The subordination of women is furthered by the practices of very early marriage for girls, a large number of children, and the expectation that girls will not complete even a secondary education. For example, one of the few sociological studies (Bennigsen and Lemercier-Quellejay, 1966) on Moslem education showed that in one major area, only two per cent of the Moslem girls finished high school. Moreover, the Minister of Education of the Uzbek Republic has reported (Shermukhamedov, 1972) that vocational training of Uzbek girls is "a very critical problem (p. 35)" and that officials are quite concerned about failures in recruiting such girls into agricultural work. In addition, among men who can afford it, it is not uncommon to have more than one wife simultaneously. Only the first goes through civil ceremonies. The others go through religious but not civil ceremonies so that charges of bigamy can be avoided. Another factor is the dearth of intermarriage between Russians and Soviet Moslems. The extent of this phenomenon is unknown but is extremely small. The intermarriage of Moslem men with Russian women is more common than Moslem women with Russian men, a manifestation chiefly of the control the Moslem family exercises over its daughters. Intermarriage is growing and Soviet sociologists and government officials are showing greater interest in it. (Aspaturian, 1968)

Very closely related is the factor of the extended family. Although still found sometimes in the purest form, its incidence has declined markedly under decades of attack by the Russians. Its significance among urban Moslems now shows generally in an unusually strong interest in genealogy. Its significance among rural Moslems lies in the extremely important fact that the brigade, currently the basic organizational unit of a collective farm, may well consist exclusively of a kinship group of 100 to several hundred persons. In such a case, decision-making authority and habits tend to be quite traditional. Personal, private income is sometimes re-allocated by the earner to a fund for the good of the kinship group. Three American behavioral scientists have noted the following (Medlin, Carpenter, and Cave, 1965):

Curiously enough, there are indications that practices carried out by the Soviets in the collectivization of farms tend to regroup the farm family in such a way as to reorganize it along traditional lines. . . . Generally, these are organized in such a way as to permit the elders to allocate work roles to members of the brigades who just happen to be members of the same extended family (pp. 129-130).

Two detailed research projects of Soviet cultural anthropologists have been reported in scholarly publications (Korbe, 1951; Snesarev, 1957), both supporting the general thesis. In one of them, the investigator found that 100 of the 116 households on one kolkhoz belonged to one extended family (Snesarev, 1957).

The Buddhist cultures of the U.S.S.R. do not exhibit such a high profile and, indeed, are little analyzed. The Buryat Institute of Social Sciences, however, is conducting some field research and gives considerable attention to Buddhist art forms and medical behavior. Buddhist medicine differs drastically from the main stream of Soviet practice (Parfionovich, Seppi, Estrin, and Yevdokimova, 1970). Research on and knowledge of Buddhist cultural tradition are hampered by the poor relations between China and the Soviet Union, for the field work necessarily involves Tibet, Sinkiang, and the several thousand miles of border between the two nations.

The prevailing attitudes of Russians and other Soviet Europeans toward the Asians serve as an extremely important backdrop for behavior of all types, including consumption. It may be said without much trepidation that the Soviet Europeans feel superior in many ways. This is exemplified in one Soviet economist's statement that appeared in five Central Asian publications recently (Tursenov, 1971):

The Uzbeks, Kazahks, Kirgiz, Tadzhiks, Turkmen, and Karakalpaks under the guidance of the Communist Party and with the aid of the fraternal Great Russian people and the rest of the peoples of the U.S.S.R. have made the great leap from backwardness to progress and from feudalism to socialism and have thus truly achieved equality in a short historical period.

The New Frontiers

Settlement of frontier areas has meant attention to two regions: (1) the North including the northern portions of European Russia and the northern portions of Siberia; and (2) the desert and mountainous districts bordering the People's Republic of China, Mongolia, and Afghanistan. These areas have had to be treated as very special markets. Of these two areas and their planned settlement the well known Russian economic geographer V. V. Pokshishevsky (1969) has said:

. . . experience shows that as soon as we relax our attention the slightest bit to the entire complex of economic, social and living conditions in regions where the national economy urgently requires a population influx, this influx, despite the efforts made in organized recruitment, the distribution of young specialists, etc., not only tapers off but actually changes to an outflow.

The North

Although the holding of northern areas was not an insignificant reason, the development of extractive industries was the more important reason for settling persons there. The settlement process has been a planned one ever since the 1917 Revolution. The settlers have come from three sources, but the number of persons obtained from each has not been announced, if indeed it is known. First, there has been forced re-location of some prisoners. Second, there has been a long-standing appeal to general idealism and specifically to the building up of the nation, except during World War II, of

course, when re-settlement of males to the frontier was highly restricted. Third, a partial overlap of the second source, there has been a program of incentives to encourage people to relocate.

The incentives have been higher pay, longer vacations, better pensions, and some privileges in housing and access to education. Officials have divided the North into two regions and graded the incentives according to hardship and relative difficulty of attracting people. Armstrong (1965) cites Russian authorities who admit that a typical construction worker in Noril'sk gets 2.4 times as much pay as he would get in central Russia. Moreover, a worker in Magadanskaya Oblast, an even more remote location, costs the state 3.5 times as much as in central Russia. Over half the workers in the north do not settle down but work awhile apparently for the high wages and subsequently leave for home.

Russian officials are hopeful that a higher fraction of people who go north will stay in the future and, to that end, are working on the improvement of living conditions. This would also increase the number of women willing to live in the north and thus encourage the formation or continuation of family life. Combinations of better living conditions with several different levels of premium pay, vacations, and pensions are possible and probably will be tried. Armstrong (1965) claims that ". . . the Soviet government has an easier task than the Canadian or United States Governments have, in so far as the drop in the standard of living which the Soviet northern settler must expect is relatively less, since his starting level is in many cases a good deal lower (p. 159)." Melnikov (1972) recently stated, however, that ". . . in the interests of speeding up the economic development of the eastern and northern regions, where immense natural resources are being increasingly exploited, it is planned in 1972-75 to introduce supplementary privileges. . . (p. 2)."

Clothing, food, and shelter require special provision in the north. Problems of clothing design were easily conquered, but late and incorrect shipments of clothing are bothersome. Despite extra efforts of officials, the inhabitants of the northern lands have a perpetual problem with diet. There is ample food, but the variety makes for boredom and, more important, deficiencies of some nutritional properties. Arnol'di and Belousova (1961) report a malnutrition rate of 11.6 per cent and a vitamin C deficiency in 80 per cent of their sample of 3,060 school children. Fruits, vegetables, and dairy products are in short supply. Considerable use of local fishery resources and minor use of reindeer for protein are developing, as are greenhouses for vegetables. In all of this, it must be remembered that the U.S.S.R. has only a tiny food canning industry and that frozen foods are still in the experimental stage there. More than anything else, the extent of settlement in the north will probably turn on the ability of local agriculture to adapt to the hostile environment and on the marketing-physical distribution systems to bring in dependably and at reasonable cost what cannot be supplied locally. Shelter has not proved to be such a formidable problem, in that the Russians early became the world's masters of permafrost construction. They routinely build adequate five-story structures on the permafrost. For the long run future they are thinking of bubble domes over communities, the same idea that has emerged in many parts of the world.

In early 1972 a highly placed Siberian public official asserted that the image of that vast area was becoming one of large industrial centers along the Trans-Siberian Railroad, but that one must not lose sight of the people living in the wilderness. (Melnikov, 1972) He stated:

When you ask people who are leaving northern places about their reasons, almost no one complains of the climate, earnings, or dissatisfaction with work, but most of them answer: "There's no place to go. Supplies are poor. It's the back country!"

It must be said in fairness that the attention paid recently to the remote regions of Siberia and the Far East has grown noticeably. Production forces have been developing at a rapid rate, the material and cultural level of the population has been rising, and trade and everyday services in regions difficult of access have been improving. . . .

But it must be admitted that supplies and services for the back country are still not so well organized.

There is a controversy between trade and transportation executives concerning the use of air freight in supplying the consumers in the north. Not enough planes are provided, many of the planes are unsuitable for perishables, and the freight rates are in dispute. The trade organizations have none of the famous Arctic adapted truck equipment for their use on the ground, although merchandise frequently must go as much as 500 miles by truck after going 500 miles on a summer-flowing river after leaving the railhead. Just as important, trade officials are unable to pay their store managers salaries comparable to those paid in other lines of work in the north and so suffer from high turnover of personnel. Also the local officials in the north do not give store managers all the privileges and fringe benefits that they give other persons, thus causing much unrest.

The Asian Borderlands

The Russian Tsars conquered much of the inner reaches of Asia, some of the current Asian borderlands, and what is known today as the Soviet Far East adjoining northeastern China and North Korea during the 17th, 18th, and 19th Centuries. For the most part the areas were sparsely populated except for a very few fertile valleys, and many of the few inhabitants were nomadic. China and Russia formulated many agreements from about 1860 to the mid-1880's dividing up this vast interior region, some of which the Chinese reject now. Similarly the Chinese repudiate some of the treaties setting the boundary lines in the Far East, as the famous Ussuri River clashes of the late 1960's exhibited.

The borderlands of interior Asia are not far from the very center of the huge region that geographer Sir Halford Mackinder called the world's "Heartland" and the "pivot of history" (Halford, 1904). His research and thought on the subject excited Western Europe's intellectuals and statesmen very much in the early 1900's and caused much more attention to focus on Russia's frontier settlement process. According to geographers of today (Peltier and Percy, 1966), Mackinder "gave emphasis to the coherence of an area of potential internal, overland communications which gave a high degree of flexibility in direction of land movement (p. 146)." This process of assimilation of the heartland has been variously interpreted, but chiefly as either strategic or economic. Economic geographer David J. M. Hooson (1964) believes that the decisive contribution of the Soviet Heartland has been economic rather than strategic. He dismisses mysticism, romanticism, and the "charmed sanctuary" thinking, and relates the borderlands primarily to economic development.

The hope for profitable trade and the attendant control of passes and oases in this rough perimeter, more than the factor of mineral industries, help explain Russia's historical interest in incorporating it into the nation and peopling it with Slavs. However, it appears to the author that pure expansionism, national pride, and political intrigue were just as important factors. And, of course, another factor was the forestalling of thoughts of independence among the indigenous inhabitants. Since World War II some importance can be attached also to the maintenance of parity with or superiority over Chinese development in the borderlands and some interest in mineral exploitation.

Both the tsars and the communist government have tried to re-settle Russians in the Asian borderlands, but in most places the population remains sparse. The Soviet Far East, adjoining northeastern China and North Korea, is resource poor. Its population is largely urbanized and the countryside quite underdeveloped. Most of the Soviet Far East population clings tenaciously to the Trans-Siberian Railroad, by which most of its supplies and output travel. One interesting aspect of this region's settlement was the experiment mainly identified with the 1930's of the Jewish Autonomous Oblast. This attempt to create a Jewish homeland as part of the U.S.S.R., called Birobidzhan, to counteract Zionism's appeal, attracted only a few Jews and subsequently failed in that mission. To this day Jews in the Soviet Union are classified by the census authorities as a separate ethnic type.

The Mongolian sector of the border appears to be clearly defined, perhaps a reflection of the fact that that country is in no condition to assert itself (Murphy, 1966). The Russians annexed the independent republic of Tannu-Tuva on the Soviet-Mongolian border in 1944 and made it an integral part of the U.S.S.R. Slavs now slightly outnumber the Tuvinians there. The southern Siberian borderlands have undergone impressive industrialization since 1917, especially since 1945, and half a dozen sizeable cities have grown up. The number of persons living in the rural areas is not great.

In the middle of the continent Soviet population drives have brought settlement within a few miles of the border of the vast Chinese province of Sinkiang. Premier Khrushchev's virgin lands policy, although questionable from a scientific viewpoint, intensified the settlement process near the Chinese border. For example, the commercial and industrial city of Alma Ata, only about 170 miles from the Sinkiang border, contains 700,000 people. The virgin lands projects also brought tens of thousands of Russians, Belorussians, and Ukrainians to the border farms. Apparently, some few persons on both sides of the border remain nomadic. The famous Turkestan-Siberian Railroad parallels the border within the U.S.S.R. and has served not only as a transport facility to carry settlers and goods but as a spine from which vertebrae of development fan out. The Russians have built a rail line from Aktogay on the Turk-Sib to the Sinkiangese border at the fabled Dzhungarian Gate, but the Chinese refuse to link up (Kueishang, 1961). Until the late 1950's most of the population of Sinkiang was Moslem and dangerously neutral, but Peking has found it possible to resettle several million persons from main stream Chinese society to that region since that time.

Not far from the Afghan border, the valleys of the Pamir Range have been settled for many centuries, primarily by Tadzhiks, but the high plateaus remain almost uninhabited. It may be recalled that, although the Soviet Union and Pakistan have no common border, the two are only eight miles apart in the Wakhan Valley district. That panhandle was created by a joint Anglo-

Russian commission in 1895 as a buffer between the British and Russian empires. The Soviet government is interested now in populating the plateaus of the Pamir.

References

- Armstrong, Terence E. Russian Settlement in the North. London: Cambridge University Press, 1965.
- Arnol'di, I.A., and Belousova, A.Z. Hygienic Problems of Acclimatization of the Far North. Moscow, 1961. Available from U.S. Department of Commerce, Office of Technical Services, Washington, D.C.
- Aspaturian, V. The Non-Russian Nationalities. In A. Kasoff (Ed.), Prospects for Soviet Society. New York: Praeger, 1968.
- Bennigsen, A., and Lemercier-Quelquejay, C. Islam in the Soviet Union. New York: Praeger, 1967.
- Brezhnev, L.I. The 24th Congress of the Communist Party of the Soviet Union: The Report of the C.P.S.U. Central Committee to the 24th Congress of the Communist Party of the Soviet Union. Pravda, March 31, 1971, 2-10.
- Bromlei, N. The Standard of Living in the U.S.S.R. Voprosy Istorii, 1966, 13.
- Chang, Kueishang. The Changing Railhead Pattern in Mainland China. Geographical Review, 1961, 51, 542.
- Hooson, David J.M. A New Soviet Heartland? New York: Van Nostrand, 1964.
- Ilyin, A. How is the Village to Grow Younger? Komsomolskaya Pravda, January 26, 1972, 2.
- Keesing, Felix M. Cultural Anthropology. New York: Holt, Rinehart and Winston, 1958.
- Korbe, O.A. The Culture, Manners and Customs of a Kazakh Kolkhoz. Sovetskaya Etnografiya, 1951, 679.
- Mackinder, Halford. The Geographical Pivot of History. Geographical Journal, 1904, 23, 421-437.
- Medlin, William K., Carpenter, F., and Cave, W.M. Education and Social Change: A Study of the Role of the School in a Technically Developing Society in Central Asia. Ann Arbor: University of Michigan, 1965.
- Melnikov, I. Back Country. Izvestia, January 15, 1972, 1-2.
- Mitrany, David. Marx Against the Peasant: A Study in Social Dogmatism. New York: Collier, 1961.
- Murphy, George G.S. Soviet Mongolia: A Study of the Oldest Political Satellite. Berkeley: University of California Press, 1966.
- Parfionovich, Yu., Seppi, I., Estrin, Yu., and Yevdokimova, N. An Ancient Culture: Dying and Living. Literaturnaya Gazeta, December 9, 1970, 12.
- Peltier, L.C., and Percy, G.E. Military Geography. New York: Van Nostrand, 1966.
- Pokshishevsky, V.V., Population Migration in the U.S.S.R. Priroda, September 1969, 67-75.

Shermukhamedov, S. The Student Chooses an Occupation. Pravda, January 18, 1972, 3.

Snesarev, G.P. Some Reasons for the Persistence of Religious and Customary Survivals Among the Uzbeks of Khorezm. Sovetskaya Etnografiya, 1957, 60.

Tursenov, Kh. The Soviet East. J.P.R.S. Editorial Report #53732, Political and Social Affairs # 166.

A CULTURAL APPROACH TO THE STUDY
OF DIFFUSION AND ADOPTION OF INNOVATIONS¹

Bernard Dubois²
Graduate School of Management
Northwestern University

Why do French consumers resist frozen foods, breakfast cereals, and hamburgers while, at the same time, to most Americans, the idea of eating horsemeat, snails and frog legs seems at least strange if not repulsive? To account for these differences in the adoption of commercial products, various types of explanation have been advanced; some have been cast in economic terms, others in physiological terms, still others in geographico-climatic terms. At a deeper level, it seems that such phenomena can be explained, at least partially in terms of differences in habits and life patterns, that is, differences in culture.

The purpose of this paper is to describe and discuss the nature of the cultural factors which affect the rate of diffusion and the rate of adoption of innovations. After having introduced the subject, the paper briefly discusses the concept of culture. Then, it analyzes how, at a macro-level, cultural norms affect the rate of diffusion of innovations. Third, at a micro-level, it presents the hypothetical relationship existing between cultural integration and individual adoption. Finally, it suggests an integrative framework which provides criteria according to which further research in the area can be stimulated and evaluated.

A certain number of examples presented in the text concern French situations. This is not intended to indicate that the scope of this paper is limited to the French culture but merely results from the greater familiarity with this particular environment by the author.

Cultural Studies in the Diffusion Literature

Although many diffusion researchers (see for example Rogers, 1962; Rogers with Shoemaker, 1971; Robertson, 1971; Katz et al., 1963 and Zaltman, 1965), recognize the importance of cultural influences upon the diffusion and adoption processes, very few attempts have been presented which concentrate upon an analysis of their functioning at a conceptual level. We know that culture is a crucial variable but still have difficulties understanding how it works. Generally speaking, social scientists (see especially Barnett, 1953; Arensberg and Niehoff, 1964; Foster, 1962), seem to have been more interested in the role of innovations as elements of cultural change than in the impact of culture upon the diffusion and adoption processes. In other words, the main stream of interest has concerned the influence of innovation upon culture rather than the influence of culture upon innovation.

Despite these qualifications, a certain number of diffusion studies exist which clearly establish the predominance of cultural factors. The three studies reported below count among the most significant of them.

In both editions of his book, Rogers (1962; 1971) relates the story of the unsuccessful diffusion of a socially desirable innovation: that of water-boiling in rural Peru. As reported by Wellin (1955) and discussed by Rogers, the reasons for the failure deeply involved the villagers' customs regarding hot, cold foods and illness; in the village under study, the basic belief system is that all foods, liquids, medicines and other materials are inherently hot or cold, irrespective of their actual temperature. Hot-cold distinctions serve essentially as a basis for discriminating between health and illness behavior. Raw water is originally considered as "cold" and therefore adequate for healthy people. Boiling it makes it less "cold" and therefore only restricted to ill persons. As a result of this, after two years of intensive efforts, only five per cent of the village population had adopted the new practice. Wellin concludes: "(Even) a practice as mundane as the domestic boiling of water is arbitrated by local cultural standards ... By virtue of their common membership in the community as such, residents of Los Molinos share many customs--including a core of common understandings about boiled and unboiled water ... This study suggests that detailed knowledge of social and cultural factors of the community is vital to the efficiency of the water-boiling program. It also suggests that useful wisdom comes not simply from knowing the scattered items of cultural belief and practice but from the appreciation that they constitute a system in which the individual parts are linked to form a meaningful structure." (1955: 100-102).

A second study of interest is the one conducted by Pedersen (1951) concerning the adoption of selected farm practices by farmers from two different cultural backgrounds. Unlike the Wellin's report, based on a case study, Pedersen's analysis is closer to a field experiment: He compared the rate of adoption of agricultural innovations concerning livestock, cropping and mechanization practices by Danish and Polish farmers settled in Central Wisconsin. All factors, except ethnic background were held constant through the selection of the two groups. In all cases, the number of Polish farmers adopting the recommended practices was much lower than the number of Danish farm operators. As explained by Pedersen, "this performance suggests that the Danish and the Polish ethnic groups constitute different universes from the standpoint of behavior or reaction to recommend dairy farming practices. The attitudes and values of the Danish group are different from the attitudes and values of the Polish group . . . The Danish farmer values education, whereas the Polish farmer places greater confidence in home training. The latter attitude tends to perpetuate the established way of performing specified operations. The orientation of the Danish farmers is out-going and community centered whereas that of the Polish farmer is family-centered. The former orientation opens up avenues of communication with outside agencies, thereby facilitating the dissemination of information in the farm community, whereas the latter tends to block the dissemination of information. Finally, the value placed upon independence and individual freedom in the Danish group tends to facilitate the acceptance of new farming practices by making a relatively complete break in operations between father and son, leaving the son free to try out new adaptations. In the Polish group, in contrast, transference of the farm from father to son is a gradual process which is tied to the training or apprenticeship of the son in farming ... In short, the culture of the Danish group facilitates the introduction of new ideas, whereas the culture of the Polish group tends to perpetuate the status quo." (1951: 48-49).

The third and last study that will be reviewed here, derived this time from a marketing environment, is a landmark in cultural research in marketing. It is Graham's study (1956) of the diffusion of five innovations: television,

canasta, supermarkets, hospital insurance and medical service insurance across six subcultural groups of contemporary America. In comparing the percentages of accepters and rejecters of each of these innovations in each of these groups, Graham found that no consistent behavior could be observed with respect to the adoption of the five innovations considered together; some innovations were accepted by some groups and rejected by others, while other innovations were accepted or rejected by all groups.

In trying to account for the differences thus revealed, Graham emphasized the cultural determinants of the adoption decision. Television for example tended to be accepted by lower class people and rejected by upper class members mainly because of a significant "penchant" for passive leisure experienced by the former subgroup. Inversely, canasta was easily adopted by upper class people given their greatest interest in active recreation and their more frequent gathering together with friends for social purposes. Graham concluded: "Of critical importance in determining the degree of acceptance of an innovation is the extent to which innovational characteristics and the culture of the receiving group are compatible. Each innovation is unique, each is compatible in different degrees with the culture of a given group. Therefore, each is accepted in different degrees by that group." (1956:99)

Although only briefly discussed here, the three examples presented above (the interested reader will find other studies dealing with the same topic in Alers-Montalvo, 1957; Apocada, 1952; Barnett, 1953; Brandner and Straus, 1954; Erasmus, 1962; Fliegel and Kivlin, 1962; Holmberg, 1952; Linton, 1936; Mead, 1955; McCorkle, 1961; Sasaki and Adair, 1952; Datt Singh, 1952; and Suttles, 1951) seem sufficient to convey the idea that culture definitely influences the degree to which innovations are accepted by individuals and spread among their groups. How this process occurs and what the mechanisms underlying it are provides the matter for the subsequent parts of this paper. However, to clarify the discussion, an initial and brief analysis of the concept of culture is in order.

The Concept of Culture: Cultural Values and Cultural Norms

Many definitions of the concept of culture have appeared in the literature (in their "Culture: A Critical Review of Concepts and Definitions," 1963, Kroeber and Kluckhohn discuss one hundred and sixty-four of them) and it is not in the scope of this paper to review and analyze them. Given our purposes, what seems more important is to understand what the essential dimensions of culture are, as far as its influence upon human behavior is concerned.

In essence, culture consists of a system of values. By values, we mean specific items, ideas or concepts which receive particular positive or negative connotation. For example, we easily recognize that in our contemporary western culture, the possession of goods is a highly valued cultural goal (Kassarjian and Robertson, 1968) while in other cultures, this is not necessarily the case (Mead, 1935). In symbolic form, we can therefore describe a given culture in terms of its "characteristic value vector" $V = (v_1, v_2, \dots, v_n)$ where $v_1, v_2, v_3, \dots, v_n$ stand for the items or ideas which are particularly valued (positively or negatively). To illustrate our point, consider the description of the American culture given by Engel, Kollat and Blackwell (1968): It is expressed in terms of an eight-element value vector; these values are: Religiosity, achievement, security, other direction, conformity, leisure, youthfulness, and urbanization. A similar description of the French culture would have probably included the

the following values: individualism, sedentarism, nationalism, social justice, freedom, logic, artistic sense and mature age. (The interested reader is referred to Brinton, 1968; Curtius, 1932; Cohen-Portheim, 1933; McKay, 1951 and Siegfried, 1930; for various presentations and discussions of the French culture.)

In order to be operational, cultural values have to be expressed in terms of norms; cultural norms refer to the tolerable patterns of overt behavior which are to be observed by the members of a given culture. It is through norms that individuals become acculturated and often, the concept of culture is defined in terms of norms rather than in terms of the values underlying them; see for example the classic definition of a culture given by Linton (1948): "A configuration of learned behavior and results of behavior whose component elements are shared and transmitted by the members of a particular society." Incidentally, it can be noticed that as long as the operationalization process which transforms values into norms is accurately understood, the decision of defining a culture in terms of norms rather than in terms of values becomes a matter of personal convenience; in particular, the description of a given culture should amount to the same thing whether it is expressed in terms of norms or in terms of values. Personally, we prefer to define a culture in terms of values only to the extent that values, more general and more abstract than norms, seem to represent more adequate vehicles to convey the essence of a culture.

As we will see later, those norms and values play a critical part in determining the diffusion and adoption rates of an innovation.

So, to sum up the discussion thus far:

(1) A given culture can be best described by identifying the particular set of values which constitutes its essence. (2) To be operational (that is, to effectively affect the behavior of its members) the values of a culture are translated into norms. (3) The thesis presented in this paper is that an analysis of these norms is of critical relevance if we are to understand the impact of culture upon the macro-process of diffusion as well as the micro-process of adoption of an innovation.

Cultural Norms and the Diffusion Process

In order to understand the impact of culture upon diffusion, it is first necessary to realize that the diffusion of an innovation never operates in a vacuum but within the boundaries of a social system, the precise nature of this social system depending upon the type of innovation being considered (for example, "in most farming studies, the social system is defined as a farming community, often a county. It is also possible to think of the social system in terms of age, income, social class or any other criteria of market segmentation") (Robertson, 1971).

Given our purposes, it is convenient to conceive of a social system as the output of a social stratification process whereby cultural values are distributed among its members. Any known society has its stratification system; it is less the amount of stratification than the rules according to which this stratification is operated which differentiate one system from another. These rules are of course the cultural values and norms which were discussed in the preceding section of this paper. To illustrate, we can take the example of certain Polynesian groups which accord high prestige to first-born children irrespective of sex. In this case, birth order is the relevant criterion. Similarly, we recognize that in France until recently, the ownership of land was traditionally synonymous

with wealth and high status. Such a system would have appeared completely "irrelevant" to the Arapesh of New Guinea who tend to believe that man belongs to the land rather than the contrary (Mead, 1935).

Among the various norms existing in a social system, one is of particular relevance given our interests: the norm which prescribes the behavior individuals will follow with respect to novelty and change. It is well known that some cultures value positively novelty and change for their own sake. The Americans for example are said to be attracted by the new. By contrast, countries like France consider change as having much less positive appeal. (Often, this attitude is linked with a basic fatalism and/or cultural ethnocentrism.) Rogers (1962, 1969) was instrumental in presenting a theoretical framework to account for these differences. He basically distinguishes two ideal types of societies: the modern one in which norms tend to systematically favor innovativeness and welcome change and the traditional one which tends to resist the new. According to Rogers, the major dimensions along which modern and traditional social systems differ are: the level of technology, the level of education, the degree of cosmopolitanism, the existence of a "rational mind" and the ability to empathize.

Although such a framework is of some value as a contribution to analysis of the impact of culture upon the diffusion of innovations, it should be kept in mind that the traditional-modern norm is but one way in which culture may exert its influence.

More generally, it is useful to recognize that an innovation is always perceived with respect to a given number of specific attributes such as relative advantage, complexity, divisibility, etc. (Rogers with Shoemaker, 1971; Zaltman and Lin, 1971). Among these attributes, the one which is of particular relevance here is compatibility. Compatibility refers to the degree of concordance existing between the given characteristics of the innovation and the central values of the socio-cultural system under consideration. Numerous studies have been conducted which demonstrate how innovations have failed to diffuse because of a lack of compatibility with the existing norms of the socio-cultural system. Foster (1962) for example reports that agricultural extension programs in Buddhist countries have encountered tremendous problems in pest control because the religious prohibition against taking life in any form was logically incompatible with the direct approach to the problem through insecticides. Similarly, people who practice polytheism do not offer great resistance towards Christian missionaries while proselyting among monotheistic people was a much more difficult task.

In France, a few years ago, an attempt was made to launch California wines on the market. Over 48,000 bottles were imported and put into sale at the major department stores in Paris. Although importers were careful enough to remove all French-like labels (champagne, burgundy, etc.) which would have irritated the French consumers, the failure was almost complete. Apparently, the proposed innovation was incompatible with a culturally-based image of wine as a "home-made" product present in the minds of most Frenchmen.

The limited success of canned soup in France can be explained in similar fashion. Canned soup is unacceptable to most rural families for whom soup is a complex and traditional dish which has to be prepared and cooked for several hours.

Recent advances in multidimensional analysis offer an interesting framework for the operationalization of the compatibility dimension. Going back to our "characteristic value vector," it can be seen that a given culture can be

graphically represented as a point (or a subspace in the case of less than interval dimensions) in a multidimensional space where the axes stand for the characteristic values expressed in bipolar form (inner direction as opposed to other direction, leisure as opposed to work and so on ...) Then, it becomes possible to locate various innovations as points in the same multidimensional space. To accomplish this, one has to obtain data on the relative standing of each innovation with respect to each characteristic value. In other words, we want to know the nature of the relationship existing between a given innovation (T.V. for example) and such well-established values as conformity, leisure, achievement, and so forth. Using this framework, it becomes possible, through the computation of "distances" lying between the various innovation points and the culture-point, treated here as an ideal point, to compare and predict diffusion rates for a selected number of innovations within the boundaries of a given culture. Although we fully recognize that the usefulness of such a framework is completely dependent upon its ability to express the main attributes of an innovation in terms of the values of the culture under consideration, we believe that it is sufficiently promising to warrant at least small-scale trial.

In the light of the discussion presented above, it appears logical to reaffirm the following generalization:

The rate of diffusion of an innovation among the members of a particular social system depends upon the nature of the relationship existing between the perceived attributes of the innovation and the central values (or norms) of the social system.

More specifically:

- (1) The rate of diffusion will be high (A high rate means a larger percentage of the social system members adopting the innovation in a relatively short period of time) when the perceived attributes of the innovation are compatible (consistent) with the central values of the social system.
- (2) The rate of diffusion will be low when the perceived attributes of the innovation are incompatible (inconsistent) with the central values of the social system.

Table 2 summarizes these hypotheses:

Table 2

Cultural Norms (Values) and the Rate of Diffusion

Level of compatibility existing between the perceived attributes of the innovation and the values of the social system		
	High	Low
Resulting Rate of Diffusion	High	Low

Cultural Integration and Individual Adoption

Having discussed how cultural norms and values affect the rate of diffusion of an innovation within a particular social system is in fact having treated one half of the problem we consider here. For our theoretical framework to be complete, we still have to account for the influence of socio-cultural factors upon the individual rate of adoption. In other words, we have to transfer our analysis from a macro-sociological to a micro-psychosocial level.

To understand the mechanism of this influence, it is crucial to recognize that any individual does not adhere to the culture of his society to the same degree. In other words, the members of a given social system do not look like the homogeneous mass that we have seemed to imply so far. Each one is an individual with his own attitudes and cognitions. Generally speaking, an individual's adherence to his culture is determined by the particular position he occupies within the socio-cultural hierarchical system of his society. Such a position is generally referred to as his social status (Linton, 1945). The specific set of behaviors an individual is expected to observe according to his status is called his role. Role is therefore the operationalization of status. If we remember our previous discussion of cultural values and norms, it is interesting to parallel the kind of analysis which is presented now. More specifically:

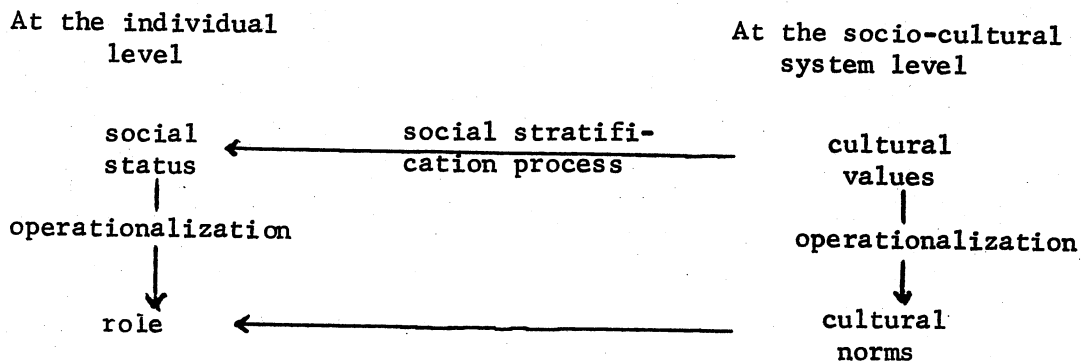


Figure 1

The influence of culture upon the individual decision of adopting or rejecting an innovation can thus be expressed at two equivalent levels: at the abstract level, it is expressed in terms of values and status; at a more empirical level, in terms of norms and role.

How does this mechanism operate? A key concept to be introduced at this level of analysis is the concept of cultural integration. Cultural integration refers to the relative degree of conformity existing between an individual's behavior and the central norms of the socio-cultural system of which he is a member. Highly integrated people strictly observe the norms existing in their culture. Poorly integrated people are sometimes referred to as "deviants" or "marginal individuals." Generally, a high social status is associated with a high degree of integration while a low social status is associated with a poor level of integration.

It should be noted that the potential usefulness of multidimensional analysis discussed in the previous section also applies here. After studying the extent to which a selected number of individuals follow the norms of their culture, it becomes possible to locate these people as points in the same

multidimensional space alluded to before and to make predictions concerning their respective adoption of an innovation that would be perfectly compatible with the culture under consideration.

An interesting illustration of how the mechanism of differential cultural integration operates is found in the experience of supermarkets introduced into France. Local organizations, social meetings and clubs are much less developed in France than in the United States. As a result, the daily shopping experience is, for the French housewife, a major way of maintaining and developing her contacts with the outside world. To this extent, supermarkets which rest on self-service and weekly shopping in bulk, are not likely to be welcomed by her. Things however are rapidly changing now in France. New generations, through modern mass media, are in constant contact with the outside world. Often working under severe time pressure, they do not have time to spend in daily shopping. In other words, they are much less respectful of the traditional norms governing the shopping experience. To them, supermarkets mean substantial savings both in time and money; therefore, they readily accept them.

According to our discussion, we can now attempt the following generalization:

The rate of individual adoption of an innovation by a member of a particular social system depends upon his level of integration within the socio-cultural system under consideration.

More specifically:

- (3) The rate of individual adoption will be high (in terms of the celerity with which the adopter goes through the stage of adoption process) when the potential adopter is deeply integrated within his socio-cultural system AND when the social values of the system favor novelty and change.
- (4) The rate of individual adoption will be low when the potential adopter is a relatively marginal member of his social system AND when the social values of the system favor novelty and change.
- (5) The rate of individual adoption will be low when the potential adopter is highly integrated within his socio-cultural system AND when the social values of the system resist novelty and change.

It is now interesting to combine the theoretical statements presented above into the form of the following matrix:

Table 3

Cultural Factors in the Diffusion
and Adoption of an Innovation

		Level of integration of the individual with respect to his social system	
		High	Low
Level of compatibility between the perceived attributes of the innovation and the values of the social system	High	High rate of Diffusion 1	Indeterminate* 3
	Low	Low rate of diffusion 2	Indeterminate* 4

* See the corresponding development in the text.

It is easily observed that the propositions presented above can be derived from an internal analysis of the cells of the matrix. When the attributes of an innovation are highly compatible with the cultural norms and values of a well integrated social system (cell 1), the innovation is likely to diffuse rapidly among its members. Inversely, if another innovation highly incompatible with these norms and values is introduced into the same social system, it will most probably be strongly rejected (cell 2).

Perhaps even more interesting are the situations presented in cells 3 and 4. In cell 3, the innovation is consistent with the social system values and norms and therefore would tend to rapidly diffuse but the level of integration is so low that its adoption remains limited. Cell 4 presents the situation of a poorly integrated social system in which an innovation that hardly matches the central values of the dominant culture is introduced. In this case, predictions are hard to make because if those individuals who strictly respect the dominant norms are likely to reject the innovation, those who do not conform might very well be tempted to accept it (especially if it questions the established status quo). This is a situation which is sometimes discussed in the diffusion literature: it is argued that non-integrated members of a social system are often more likely to adopt innovations than highly-integrated people mainly because the former group has "nothing to lose" in making such a decision. If the anticipated consequences of the adoption of the innovation are such that a restructuring of the existing social system is likely to take place, they might even have something to gain. In terms of our matrix, we understand that this situation can only happen when the cultural values or norms of the system under consideration do not favor novelty and change. If, conversely, the system is of a more "modern" type, it is highly improbable that well integrated members will not attempt to adopt the innovation under consideration, given the high social rewards attached to such a decision.

Other similar analyses could be derived from such a matrix. In fact, we hope that it provides a useful framework for analyzing and interpreting research results as well as suggesting new avenues for further investigation. In particular, it is interesting to note that the cultural approach to the study of the diffusion of innovations which is considered here integrates dimensions such as "centrality" or "compatibility" which have tended to be separately presented in the diffusion literature. To a lesser extent, this approach also sheds new light upon dimensions such as social cost, terminality, gatewayability and ego-involvement (Zaltman and Lin, 1971).

Clearly, empirical evidence is needed to support the theoretical framework presented above. Within the existing tradition of diffusion studies, such a validation is hardly possible. The reason is that diffusion researchers have so far focused upon the diffusion of an innovation within a particular social system without attempting to control for the specificities of the social system under study. In other words, what seems to be needed is an inter- rather than intra-cultural approach to diffusion research. If adopted, this orientation could lead, we believe, to major contributions to diffusion and adoption theories.

Conclusion

As mentioned in the introductory section of this paper, the relationship between culture and diffusion and adoption of innovations is two-fold. Because we felt that this side of the coin has been relatively neglected in the diffusion literature, we deliberately decided to concentrate our attention upon the

processes whereby cultural factors affect the diffusion and adoption of innovations. Clearly, this orientation does not mean that we do not consider innovations as determinant elements of cultural change. In fact the relationship existing between culture and innovation is a deeply interactive one: As we have attempted to indicate in this paper, the rate of diffusion and the rate of adoption of an innovation within a socio-cultural system is critically dependent upon the nature of the cultural values and norms underlying the system. Conversely, the consequences of the acceptance of an innovation can be such that some aspects of the cultural system in which the innovation takes place, may be affected. In such cases, a cultural change movement is generated. The innovation contributes to the development of a cultural system rather than being affected by it.

Interestingly and paradoxically enough, we may hypothesize that the innovations which are accepted with the least degree of resistance because of a high compatibility of their attributes with the cultural norms of a well integrated system for example, are likely to have the less disruptive effects, while highly "discontinuous" innovations will in the long run and if adopted offer the best opportunities for cultural change.

Footnotes

1. The author is indebted to Professor G. Zaltman and Mr. C. Pinson for their critical comments on an earlier draft and their helpful suggestions.
2. Bernard Dubois is a doctoral candidate in marketing at the Graduate School of Management, Northwestern University.

References

- Alers-Montalvo, M. Cultural change in a Costa Rican village. Human Organization, 1957, 15, 2-7.
- Apocada, A. Corn and custom: introduction of hybrid corn to Spanish American farmers in New Mexico. In E. H. Spicer (Ed.), Human problems in technological change. New York: Russell Sage Fdn, 1952.
- Arensberg, C. M. and A. H. Niehoff. Introducing social change. Chicago: Aldine, 1964.
- Barnett, H. G. Innovation: the basis of cultural change. New York: McGraw-Hill, 1953.
- Brandner, L. and M. A. Straus. Congruence versus profitability in the diffusion of hybrid sorghum. Rural Sociology, 1959, 24, 381-383.
- Brinton, C. The Americans and the French. Cambridge, Mass.: Harvard University Press, 1968.
- Cohen-Portheim, P. The spirit of France. New York: E. P. Dutton, 1933.
- Curtius, E. R. The civilization of France. New York: MacMillan, 1932.
- Datt Singh, R. The village level: an introduction of green manuring in rural India. In E. H. Spicer (Ed.), Human problems in technological change. New York: Russell Sage Fdn, 1952, 55-67.
- Engel, J. F., D. T. Kollat and R. D. Blackwell. Consumer behavior. New York: Holt, Rinehart and Winston, 1968.
- Erasmus, C. J. Introducing new agricultural practices in Latin America. Migration News, 1962, 11, 7-12.
- Fliegel, F. C. and J. E. Kivlin. Farm practice attributes and adoption rates. Social Forces, 1962, 40, 364-370.
- Foster, G. M. Traditional cultures and the impact of technological change. New York: Harper, 1962.

- Graham, L. S. Class and conservatism in the adoption of innovations. Human Relations, 1956, 9, 91-100. Also see L. S. Graham, Cultural compatibility in the adoption of television. Social Forces, 1954, 33, 166-170.
- Holmberg, A. R. The wells that failed: an attempt to establish a stable water supply in Viru Valley, Peru. In E. H. Spicer (Ed.), Human problems in technological change. New York: Russell Sage Fdn, 1952, 113-123.
- Katz, E., M. L. Levin and H. Hamilton. Traditions of research on the diffusion of innovations. American Sociological Review, 1963, 28, 237-252.
- Kassarjian, H. H. and T. S. Robertson. Perspectives in consumer behavior. Glenview, Ill.: Scott, Foresman & Co., 1968, 409.
- Kroeber, A. L. and C. Kluckhohn. Culture: a critical review of concepts and definitions. Papers of the Peabody Museum, 1952, 47.
- Linton, R. The study of man. New York: Appleton-Century-Crofts, 1936. See especially chapter 11.
- Linton, R. The cultural background of personality. New York: Appleton-Century Cy Inc., 1945.
- McCorkle, T. Chiropractic: a deviant theory of disease and treatment in contemporary western culture. Human Organization, 1961, 20, 20-22.
- McKay, D. The United States and France. Cambridge, Mass.: Harvard University Press, 1951.
- Mead, M. Sex and temperament in three primitive societies. New York: William Morrow & Co., 1935
- Mead, M. Cultural patterns and technical change. New York: New American Library, 1955.
- Pedersen, H. A. Cultural differences in the acceptance of recommended practices. Rural Sociology, 1951, 16, 37-49.
- Robertson, T. S. Innovative behavior and communication. New York: Holt, Rinehart and Winston, 1971.
- Rogers, E. M. Diffusion of innovations. New York: The Free Press, 1962.
- Rogers, E. M. with F. F. Shoemaker. Communication of innovations. New York: The Free Press, 1971.
- Rogers, E. M. with L. Svenning. Modernization among peasants: the impact of communication. New York: Holt, Rinehart and Winston, 1969.
- Sasaki, T. and J. Adair. New land to farm: agricultural practices among the Navaho Indians of New Mexico. In E. H. Spicer (Ed.), Human problems in technological change. New York: Russell Sage Fdn, 1952, 97-111.
- Siegfried, A. France: a study in nationality. New Haven: Yale University Press, 1930.
- Wellin, E. Water boiling in a Peruvian town. In B. D. Paul (Ed.), Health, culture and community. New York: Russell Sage Fdn, 1955.
- Zaltman, G. Marketing: contributions from the behavioral sciences. New York: Harcourt, Brace and World, 1965.
- Zaltman, G. and N. Lin. On the nature of innovations. American Behavioral Scientist, 1971, 14, 651-673.

AUTHOR INDEX

Ahmed, S.A.	34	Kaplan, L.B.	382
Anderson, B.B.	436	Keiser, S.K.	602
Anderson, R.E.	67	Kendall, C.L.	349
Andrews, R.B.	679	Kinnear, T.C.	34
Angelmar, R.	586	Kollat, D.T.	576
		Kuehl, P.G.	602
Beery, A.	446	Kunreuther, H.	660
Bettman, J.R.	394, 679	Kushner, R.I.	546
Beuttenmuller, M.	428		
Bither, S.W.	9	Landon, E.L., Jr.	1, 213
Blackwell, R.D.	576	Lehmann, D.R.	526
Blattberg, R.	240	Lingoes, J.C.	689
Brund, A.V.	456	Lutz, R.J.	800, 812
Burnkrant, R.E.	807		
		Mazis, M.B.	417, 428
Calder, B.J.	812	Mittelstaedt, R.A.	101
Clawson, C.J.	522	Moinpour, R.	341
Clemhout, S.	107		
Cousineau, A.	796	Nakanishi, M.	61
		Olson, J.C.	167
Darden, W.R.	475	Pessemier, E.A.	456
Davidson, J.D.	294	Pfaff, A.	713
Dawson, L.E., Jr.	536	Pfaff, M.	689
Deering, B.J.	145, 404	Philpot, J.C.	201
Dolich, I.J.	9	Pinson, C.	586
Dubois, B.	840	Pollay, R.W.	594
		Pratt, R.W., Jr.	742
Efron, D.C.	759		
Engel, J.F.	576	Rao, V.R.	125
		Reizenstein, R.C.	201
Fiedler, J.A.	279	Reynolds, F.D.	475
Flaschner, A.B.	536	Roscoe, A.M., Jr.	258
Friedman, M.P.	361	Russ, F.A.	349
Gardner, D.M.	20	Scott, J.E.	213
Granbois, D.H.	502	Sen, S.	240
Green, P.E.	304	Settle, R.B.	417
Green, T.V.	830	Sheth, J.N.	258, 562
		Sims, J.T.	80
Haefner, J.E.	356	Stokes, R.C.	187
Haines, G.H., Jr.	759	Suh, R.H.	512
Hair, J.F., Jr.	67	Summers, J.D.	502
Handy, C.H.	738	Sweeney, D.J.	201
Hansen, F.	220	Szybillo, G.J.	180
Heimbach, J.T.	746		
Hemple, D.J.	816	Taylor, J.R.	34
Howard, J.A.	96	Taylor, R.H.	213
Hunt, H.K.	370	Thorelli, H.B.	192
Jacoby, J.	145, 167, 180, 382, 404, 632, 746	Weinreich, R.P.	325
Jain, A.K.	304		

(continued on next page)

Wiley, J.B.	341
Wilkie, W.L.....	325
Willenborg, J.F.....	783
Williams, F.L.....	490
Wind, Y.....	304
Winter, F.W.....	790
Wright, P.L.....	316, 796
Woodside, A.G.	650
Zaltman, G.....	586

ACR ADVISORY COUNCIL FOR 1972

Mary Gardner Jones	Federal Trade Commission
Barbara M. Burns	Deputy Assistant Secretary for Consumer Affairs, HEW
Alvin A. Achenbaum	J. Walter Thompson
Stanley D. Shores	Proctor & Gamble
Fred I. Posner	General Foods
Jacob Jacoby	Purdue University
Gordon E. Bivens	University of Missouri
Paul E. Green	University of Pennsylvania
James F. Engel	Wheaton College
Mary Ellen Simon	Food and Drug Administration, HEW

OFFICERS OF THE ASSOCIATION (1973)

President	Robert W. Pratt, Jr., General Electric Company
Vice President and President Elect	William D. Wells, University of Chicago
Treasurer	David M. Gardner, University of Illinois
Director	Harold H. Kassarjian, UCLA
Director	Barbara M. Burns, HEW
Executive Secretary	Philip F. Kuehl, University of Maryland
Publications Committee	Jacoby Jacoby, Purdue University
Conference Committee (1973) Chairman (and Editor of 1973 proceedings)	Scott Ward, Harvard University
Membership Committee	James Stafford, University of Houston

